

2

AIR FORCE



AD-A208 245

S DTIC
ELECTE
MAY 30 1989
D

HUMAN RESOURCES

**DEVELOPMENT OF ARMY JOB KNOWLEDGE TESTS
FOR THREE AIR FORCE SPECIALTIES**

Barbara A. Bentley
Kathleen L. Ringenbach
James W. Augustin

Universal Energy Systems, Incorporated
Manpower, Personnel, and Training Division
4401 Dayton-Xenia Road
Dayton, Ohio 45432-1894

TRAINING SYSTEMS DIVISION
Brooks Air Force Base, Texas 78235-5601

May 1989

Interim Technical Paper for Period February - December 1987

Approved for public release; distribution is unlimited.

LABORATORY

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235-5601

008

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0182

1a. REPORT SECURITY CLASSIFICATION Unclassified		1b. RESTRICTIVE MARKINGS										
2a. SECURITY CLASSIFICATION AUTHORITY		3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.										
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE												
4. PERFORMING ORGANIZATION REPORT NUMBER(S)		5. MONITORING ORGANIZATION REPORT NUMBER(S) AFHRL-TP-88-11										
6a. NAME OF PERFORMING ORGANIZATION Universal Energy Systems, Incorporated	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Training Systems Division										
6c. ADDRESS (City, State, and ZIP Code) 4401 Dayton-Xenia Road Dayton, Ohio 45432-1894		7b. ADDRESS (City, State, and ZIP Code) Air Force Human Resources Laboratory Brooks Air Force Base, Texas 78235-5601										
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Air Force Human Resources Laboratory	8b. OFFICE SYMBOL (If applicable) HQ AFHRL	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F41689-86-D-0052/SB5861-0432										
8c. ADDRESS (City, State, and ZIP Code) Brooks Air Force Base, Texas 78235-5601		10. SOURCE OF FUNDING NUMBERS <table border="1"><tr><td>PROGRAM ELEMENT NO. 62703F</td><td>PROJECT NO. 7734</td><td>TASK NO. 13</td><td>WORK UNIT ACCESSION NO. 01</td></tr></table>		PROGRAM ELEMENT NO. 62703F	PROJECT NO. 7734	TASK NO. 13	WORK UNIT ACCESSION NO. 01					
PROGRAM ELEMENT NO. 62703F	PROJECT NO. 7734	TASK NO. 13	WORK UNIT ACCESSION NO. 01									
11. TITLE (Include Security Classification) Development of Army Job Knowledge Tests for Three Air Force Specialties												
12. PERSONAL AUTHOR(S) Bentley, B.A.; Ringenbach, K.L.; Augustin, J.W.												
13a. TYPE OF REPORT Interim	13b. TIME COVERED FROM Feb 87 TO Dec 87	14. DATE OF REPORT (Year, Month, Day) May 1989	15. PAGE COUNT 56									
16. SUPPLEMENTARY NOTATION												
17. COSATI CODES <table border="1"><tr><th>FIELD</th><th>GROUP</th><th>SUB-GROUP</th></tr><tr><td>05</td><td>08</td><td></td></tr><tr><td>05</td><td>09</td><td></td></tr></table>		FIELD	GROUP	SUB-GROUP	05	08		05	09		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number) job knowledge test, performance measurement, multiple-choice test, technology transfer. (SDW)	
FIELD	GROUP	SUB-GROUP										
05	08											
05	09											
19. ABSTRACT (Continue on reverse if necessary and identify by block number) This research and development (R&D) effort was performed under the Joint-Service Job Performance Measurement Project in conjunction with the development of Walk-Through Performance Tests. This report documents the development and pre-test of task-based multiple-choice job knowledge tests for three Air Force specialties, using methodologies developed by the Army. This R&D was initiated to determine whether measurement technology developed in one Service could be effectively transferred to another Service. The resulting tests and data suggest a strong similarity to those developed by the Army. A thorough review of data collected during large-scale data collection will provide a basis for making more definitive observations concerning the accomplishments of the present research effort. <i>Keywords:</i>												
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT <input checked="" type="checkbox"/> UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS RPT <input type="checkbox"/> DTIC USERS		21. ABSTRACT SECURITY CLASSIFICATION Unclassified										
22a. NAME OF RESPONSIBLE INDIVIDUAL Nancy J. Allin, Chief, STINFO Branch		22b. TELEPHONE (Include Area Code) (512) 536-3877	22c. OFFICE SYMBOL AFHRL/SCV									

DD Form 1473, JUN 86

Previous editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE

Unclassified

**DEVELOPMENT OF ARMY JOB KNOWLEDGE TESTS FOR
THREE AIR FORCE SPECIALTIES**

**Barbara A. Bentley
Kathleen L. Ringenbach
James W. Augustin**

**Universal Energy Systems, Incorporated
Manpower, Personnel, and Training Division
4401 Dayton-Xenia Road
Dayton, Ohio 45432-1894**

**TRAINING SYSTEMS DIVISION
Brooks Air Force Base, Texas 78235-5601**

Reviewed and submitted for publication by

**Hendrick W. Ruck
Technical Advisor
Training Systems Division**



Accession For	
NTIS GRA&I	<input checked="checked" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

This publication is primarily a working paper. It is published solely to document work performed.

SUMMARY

This document describes the development and pretest of task-based multiple-choice job knowledge tests for three Air Force specialties. These tests were developed using Army methodology as part of the Joint-Service Job Performance Measurement Project. These preliminary results suggest that the technology developed by the Army for development of multiple-choice job knowledge tests was successfully transferred from one Service to another.

PREFACE

In the early 1980's, the Office of the Secretary of Defense and the military Services initiated a research and development (R&D) effort designed to develop a technology for measuring on-the-job performance. All four Services were tasked to develop different types of performance measures. In order to reduce developmental time and expense, the Services are investigating the possibility of transferring developmental technologies from one Service to another. The Air Force has been designated the lead Service of this "transfer-of-technology" R&D effort. This report documents the transfer of Army methodologies in the development of task-based job knowledge tests for three Air Force specialties.

TABLE OF CONTENTS

	Page
I. INTRODUCTION	1
Background	1
Objective	1
II. GENERAL PROCEDURES	2
Review of Army Methodology	2
Preparation of Test Development Strategy	3
III. DEVELOPMENT OF THE AEROSPACE GROUND EQUIPMENT MECHANIC (AFS 423X5)	
JOB KNOWLEDGE TEST	3
The AGE Walk-Through Performance Test (WTPT)	3
JKT Item Development	4
Pilot Test	5
Item Revision	6
Pretest	6
Pretest Data Analysis	8
Final Revisions	9
IV. DEVELOPMENT OF THE AIRCREW LIFE SUPPORT SPECIALIST (AFS 122X0)	
JOB KNOWLEDGE TEST	9
The Aircrew Walk-Through Performance Test (WTPT)	9
JKT Item Development	12
Pilot Test	12
Item Revision	13
Pretest	13
Pretest Data Analysis	14
V. DEVELOPMENT OF THE PERSONNEL SPECIALIST (AFS 732X0)	
JOB KNOWLEDGE TEST	18
The Personnel Walk-Through Performance Test (WTPT)	18
JKT Item Development	18
Pilot Test	21
Item Revision	23
Pretest	24
Pretest Data Analysis	24
VI. CONCLUSION	25
REFERENCES	26

Table of Contents (Concluded)

	Page
APPENDIX A: AEROSPACE GROUND EQUIPMENT PRETEST AND FINAL TEST CHARACTERISTICS	27
APPENDIX B: ITEM ANALYSIS FOR PRETEST VERSION OF THE AGE (AFS 423X5) JOB KNOWLEDGE TEST	28
APPENDIX C: AIRCREW LIFE SUPPORT PRETEST AND FINAL TEST CHARACTERISTICS	34
APPENDIX D: ITEM ANALYSIS FOR PRETEST VERSION OF THE AIRCREW LIFE SUPPORT (AFS 122X0) JOB KNOWLEDGE TEST	35
APPENDIX E: ITEM ANALYSIS FOR PRETEST VERSION OF THE AIRCREW LIFE SUPPORT (AFS 122X0) MAC, SAC, AND TAC JOB KNOWLEDGE TEST	38
APPENDIX F: PERSONNEL JKT CHARACTERISTICS	44
APPENDIX G: ITEM ANALYSIS FOR PRETEST VERSION OF THE PHASE I PERSONNEL SPECIALTY (AFS 732X0) JOB KNOWLEDGE TEST	45

LIST OF TABLES

Table	Page
1 Aerospace Ground Equipment Mechanic (AFS 423X5) Walk-Through Performance Test Tasks	4
2 Number of Items for AGE Task Tests	7
3 Frequency of AGE (AFS 423X5) JKT Total Test Scores	9
4 Task Test Analysis for Pretest and Revised Final Version: Aerospace Ground Equipment Mechanic (AFS 423X5) Job Knowledge Test	10
5 Aircrew Life Support Specialist (AFS 122X0) Walk-Through Performance Test Tasks	11
6 Number of Items for Aircrew Life Support Specialist Task Tests	14
7 Task Test Analysis for Pretest and Revised Final Versions: Aircrew Life Support Specialist (AFS 122X0) Job Knowledge Test: Phase I	16
8 Task Test Analysis for Pretest and Revised Final Versions: Aircrew Life Support Specialist (AFS 122X0) Job Knowledge Test: Phase II	16
9 Frequency of Aircrew Life Support Specialist (AFS 122X0) Phase I JKT Total Test Scores	17
10 Frequency, Mean, and Standard Deviation of Aircrew Life Support Specialist (AFS 122X0) JKT Total Test Scores	17

List of Tables (Concluded)

Table	Page
11 Personnel Specialist (AFS 732X0) WTPT Tasks	19
12 Number of Items for Personnel Specialist Task Tests	22
13 Task Test Analysis for Pretest and Revised Final Versions: Personnel Specialist (AFS 732X0) Job Knowledge Test: Phase I	25
14 Frequency of Personnel Specialist (AFS 732X0) JKT Phase I Total Test Scores	25

DEVELOPMENT OF ARMY JOB KNOWLEDGE TESTS FOR THREE AIR FORCE SPECIALTIES

I. INTRODUCTION

Background

The Services are presently faced with the issue of establishing a link between enlistment standards and on-the-job performance. In July 1980, the Office of the Secretary of Defense and the military Services initiated a research and development (R&D) effort designed to develop a technology for measuring on-the-job performance. The Joint-Service Job Performance Measurement (JPM) Project was established to develop prototype methodologies for the measurement of job performance. In addition, a Joint-Service Working Group was established to facilitate cross-Service cooperation.

All four Services were tasked to develop hands-on methods for measuring job performance. The Air Force developed a work sample technique called Walk-Through Performance Testing (WTPT). The WTPT is a specialty-specific, task-oriented measure which combines the observation of hands-on performance and incumbent interviews for a representative set of tasks drawn from a larger pool of tasks performed by first-term airmen in a specialty. The test is administered to one examinee at a time and takes 6 to 8 hours to complete.

The four Services are also developing an array of cost effective measures as potential surrogates for the hands-on measures. Each of the Services is focusing on developing different surrogates in order to reduce the time required to develop a comprehensive performance measurement program. This approach also avoids duplication of effort among the Services. The Air Force's emphasis has been on interview testing methods and performance rating forms. The Navy's research has concentrated on the use of simulator and training device performance tests and symbolic substitute tests. The Marine Corps has investigated the use of peripheral data collection such as technical training school scores. The Army has developed Army-wide ratings of effectiveness (e.g., discipline, effort/initiative, physical fitness) and paper-and-pencil job knowledge tests.

In order to further reduce developmental time and expense, the Services are investigating the possibility of transferring from one Service to another their technology for developing surrogates. The Air Force has been designated lead Service for this "transfer-of-technology" R&D effort. The Air Force has, therefore, been tasked with attempting to replicate procedures developed by the Army for constructing paper-and-pencil job knowledge tests (JKTs) to construct JKTs for the Air Force.

Objective

The objective of the present effort was to develop paper-and-pencil job knowledge tests for three Air Force specialties (AFSs) using procedures developed and employed by the Army. The three Air Force specialties were: Aerospace Ground Equipment Mechanic (AFS 423X5), Aircrew Life Support Specialist (AFS 122X0), and Personnel Specialist (AFS 732X0). Air Force Job Knowledge Test (JKT) items were created with the goal of achieving a one-to-one correspondence to tasks included in WTPTs for these specialties. This was done in order to facilitate successful integration of newly developed JKTs into the previously established Job Performance Measurement System (JPMS).

II. GENERAL PROCEDURES

Review of Army Methodology

This research effort commenced with a thorough review of procedures followed by the Army in developing job knowledge tests (JKTs). The review began with a careful reading of the Army Research Institute (ARI) report *Development and Field Test of Task-Based MOS-Specific Criterion Measures* (Campbell, Campbell, Rumsey, & Edwards, 1985). Due to the detailed level of procedural information needed in order to duplicate Army procedures, some topics discussed in the report required elaboration or clarification. These information needs were compiled into a comprehensive set of questions.

Air Force Human Resources Laboratory (AFHRL) and contractor personnel then met in Alexandria, Virginia, with the representatives from ARI and Human Resources Research Organization (HumRRO) involved in developing Army JKTs. In addition to this formal meeting, several key Army JKT authors were contacted by telephone on numerous occasions to ascertain details on the Army project which were otherwise unobtainable.

Representatives from ARI also provided AFHRL contractor personnel with samples of documents used during test development: Soldiers Manual of Common Tasks Skill Level 1 (1982) and Soldiers Manual for 71L, Skill Levels 1/2, Administrative Specialist (1985). The JKTs for military occupational specialty (MOS) 71L, Administrative Specialist, and MOS 95B, Military Police, were also provided for review.

After careful study of the Army test development methodology, it became apparent that several aspects could not be duplicated due to differences between the Army and Air Force JPM projects. First, Army test development personnel participated in task selection and task analysis prior to item writing for hands-on performance tests and paper-and-pencil JKTs. The Air Force, however, had completed the task. Thus, Army test developers were already familiar with tasks at the beginning of item writing; but Air Force developers had to spend additional time familiarizing themselves with the various tasks that composed the WTPT.

A second difference was that the Army contractors had experts on their staff from each career field, who could be consulted throughout item development. Because the Air Force contractors did not have this convenience, provisions had to be made to obtain input from subject-matter experts (SMEs) from the target Air Force specialties during item development.

The type and number of tasks selected for test development was a third difference between the Services. The Army tests were developed to include tasks unique to an MOS as well as tasks commonly performed by all soldiers across all MOSs. The Air Force, however, concentrated on tasks unique to an AFS, with no common tasks across AFSs. This orientation resulted in Air Force tasks which tended to be somewhat technical and/or complex. In addition, the Army, as a rule, selected 30 tasks for each MOS, whereas the number of tasks selected for an AFS varied from 13 to 26. It is believed that these numbers are comparable in view of the fact that the Air Force did not select tasks performed Air Force-wide.

The fourth Army procedure that the Air Force could not duplicate was the Army's method of assigning tasks to test type. The Army simultaneously developed the hands-on and job knowledge tests for each MOS and assigned tasks to the test mode for which the task was best suited. The Air Force, on the other hand, sought to develop JKT items for every task in the previously developed WTPTs.

The fifth difference dealt with required test format. The Army used an answer sheet specifically designed to correspond with their test. The items associated with each task were always numbered beginning with number one in order to identify the beginning of a new task. The Air Force used a previously designed, machine-scorable answer sheet which required that items be numbered consecutively from task to task.

These five departures from Army procedures were not expected to have a substantial impact on the content or psychometric properties of the resulting job knowledge tests. A more detailed comparison of the Air Force and Army methodologies can be found in Augustin, Bentley, and Ringenbach (1987a).

Preparation of Test Development Strategy

After a thorough review of all Army procedures, a strategy was developed for preparing a JKT for each AFS. The strategy included the rules that would be followed when drafting, reviewing, and pilot-testing items and a method for prioritizing tasks for item development should time limitations preclude the development of items for every task.

In order to clarify JKT development requirements, it is necessary to explain briefly the WTPT task selection strategy. Individuals in each AFS perform tasks that are the same across the entire specialty and tasks that are unique to their functional area. This functional area might be a workcenter (e.g., Outbound Assignments or Classification and Training) or a major command (e.g., SAC, TAC, or MAC). Tasks performed by everyone in the career field are included in a Phase I test. A Phase II test is constructed for functional areas, with homogeneous clusters of tasks performed by first-term airmen assigned to that area. Thus, each Phase II test for an AFS contains tasks unique to a selected functional area. Each WTPT examinee is administered the Phase I test and the Phase II test corresponding to his/her functional area. Because of the homogeneity of the AGE career field, only a Phase I test was constructed; however, both the Aircrew Life Support and Personnel specialties have Phase I and Phase II tests.

As a part of the JKT development strategy, items were to be developed first for all Phase I tasks within each specialty, then for those Phase II tasks which met all of the following criteria: (a) the corresponding hands-on test discriminated between good and poor performers; (b) the task was one for which both an interview and a hands-on test had been developed; and (c) the task was included in the technical training school curriculum and noted in the Plan of Instruction (POI). Next, Phase II tasks which met two of the three criteria were to be addressed. Finally, items would be developed for tasks meeting only one or none of the criteria. Following this prioritization scheme, the test developers were able to develop items for all tasks in all three AFSs. A list of the prioritized tasks for each specialty can be found in Augustin, Bentley, and Ringenbach (1987b, 1987c) and Augustin, Ringenbach, and Bentley (1987).

III. DEVELOPMENT OF THE AEROSPACE GROUND EQUIPMENT MECHANIC (AFS 423X5) JOB KNOWLEDGE TEST

The construction of the Aerospace Ground Equipment Mechanic (AGE) Job Knowledge Test (JKT) took approximately 6 months and was a cooperative effort between three contractor personnel who were simultaneously involved in developing the Personnel Specialist and the Aircrew Life Support Specialist JKTs.

The general procedures for developing the AGE JKT, as well as the JKTs for the other two AFSs, were: (a) assessment of the WTPT, (b) JKT item development including two SME workshops, (c) pilot test, (d) item revision, (e) pretest, (f) pretest data analysis, and (g) final revisions.

The AGE Walk-Through Performance Test (WTPT)

The development of the AGE JKT began with a detailed assessment of the AGE WTPT. The AGE WTPT consists of 31 tasks and requires approximately 8 hours 15 minutes to complete. Of these 31 tasks, 16 are hands-on tasks which require the examinee to actually perform a task using available equipment, tools, directives, and any other aids. The remaining 15 are interview tasks during which the examinee explains how the task should be performed. Five of the tasks are overlap tasks: i.e., both hands-on and interview

tests were written for the same task. In these instances the content of the two task tests is so similar that JKT items were written for the hands-on version only. Thus, JKT items were written for a total of 26 tasks. A listing of those tasks appears in Table 1.

**Table 1. Aerospace Ground Equipment Mechanic (AFS 423X5)
Walk-Through Performance Test Tasks**

Task number	Task description
H260	Clean motor and generator components.
H264 ^a	Isolate engine, motor, or generator malfunctions.
I555	Prepare AGE for mobility and training exercises.
I488	Remove or install AGE tire, tube and wheel assemblies.
I286	Remove or install engine fuel pumps.
I255	Change generators and alternators.
H215	Perform electrical systems operational checks.
H154	Perform an aircraft support generator service inspection.
H251 ^a	Adjust turbine engine fuel system components.
H209 ^a	Measure resistance in AGE electrical systems.
H179	Perform a gas turbine compressor periodic inspection.
I120	Make entries on supply issue and turn-in forms.
I322	Isolate heater system malfunctions.
I340	Remove burner control valves on AGE heaters.
H284	Remove and replace engine fan belts.
I275	Remove or install carburetors.
H446	Isolate pneumatic system malfunctions.
H503	Research TOs for chassis, enclosures and drive maintenance information.
H155	Perform a service inspection on a load bank.
H162	Perform a service inspection on a hydraulic test stand.
H421 ^a	Remove or install hydraulic lines.
I181	Perform a hydraulic test stand periodic inspection.
I477	Repack wheel bearings.
H549	Inspect vehicles for safety of operation.
H300 ^a	Remove or install fuel lines and fittings.
H238	Splice electrical systems wiring.

Note. Task numbers correspond to the Occupational Survey tasks from which WTPT items were selected. H indicates that the task is a hands-on task in the WTPT; I indicates that the task is an interview task in the WTPT.

^aIndicates those hands-on tasks for which an interview task was also developed in the WTPT.

JKT Item Development

Item development began with a workshop attended by nine AGE SMEs. The SMEs jointly reviewed the WTPT tasks to identify those steps judged to be key elements within each task. A key element was defined as any step within a task having serious repercussions if not performed or if performed incorrectly. When time permitted, SMEs also identified plausible incorrect methods for performing those steps identified as key elements. These would form the basis for developing distractors (i.e., incorrect alternatives) for the items.

Following this workshop, items were drafted to assess the key elements in each task. The developmental procedure took into account all information learned during the workshop, as well as information obtained

from WTPT task analysis, Technical Orders (TOs), Air Force Regulations (AFRs), and Career Development Courses (CDCs). After items were developed representing the key elements, additional items were constructed to achieve adequate task coverage.

Items were developed to concentrate on the application of the knowledge needed to perform a task. Therefore, emphasis was placed on developing performance items in which the examinee would have to actually perform the task (e.g., reading a technical order) to identify the correct alternative. Due to the nature of the WTPT tasks, however, this was not always possible. Many of the items had to be simply performance-related, involving questions such as how the task or step should be performed.

Certain rules consistent with the Army JKT development methodology were followed during item construction. All items were written in a multiple-choice format with two to five alternatives per item. The number of alternatives was directly related to the number of plausible incorrect responses. Alternatives were written such that only one was correct. Item stems were usually limited to two lines and were worded, whenever possible, such that they could be answered without reference to the choices. Illustrations were used as much as possible to convey information in a more efficient and effective format and in a manner similar to actual job performance. Emphasis was placed on using terms which were commonly used on the job. Thus, the use of uncommon technical terms was avoided in an attempt to ensure that examinees were being tested on their job performance-based knowledge and not on their reading level. Items were also carefully examined to eliminate any inter-item cueing.

After completion of this initial item development stage, two groups of three SMEs and a test developer participated in a second workshop in which all items were reviewed by both groups for technical accuracy and correct wording. SMEs also confirmed correct answers, determined whether the distractors were plausible, and generated additional distractors, if necessary. The SMEs examined the task test illustrations for accuracy and also considered whether any items could be improved by using illustrations. Finally, the SMEs determined whether each task was sufficiently covered by the items written or if additional items needed to be developed.

Following this second workshop, a first draft of the test was composed, taking into account all the information obtained from the two groups of SMEs. Each test item was then reviewed by two contractor personnel not directly involved in writing the AGE JKT per se. One of the two reviewers was the principal investigator on the project; the other was a research assistant who was simultaneously writing JKT items for the Aircrew Life Support Specialty. The reviewers looked for consistency with item writing guidelines and inspected the items for proper spelling, grammar, and readability. In addition, the reviewers assessed whether or not the WTPT content had been adequately covered by the paper-and-pencil JKT items. Finally, the reviewers edited the items to minimize any idiosyncracies in the writing style of the principal test developer.

Following the internal review, items were arranged within each task test to reflect the order in which the individual steps of the WTPT were performed. These task tests were grouped based on the equipment used to perform the corresponding WTPT. The task tests within each equipment grouping were then ordered from least to most difficult based on difficulty information gathered during WTPT research. Although the Army arranged the task tests randomly, the task tests in the Air Force JKT were arranged from least to most difficult in accordance with standard test writing procedures. The test prepared for pilot test contained 213 items in 26 task tests.

Pilot Test

This step in the item development process did not serve as a standard or traditional pilot test, but rather as a final review and validation of the items. The AGE JKT was pilot-tested at Bergstrom AFB, Texas. The JKT was first administered to a group of five AGE specialists who each had less than 48 months in the Air Force, but had been on the job at least 6 months. This part of the pilot test was defined as the incumbent

pilot test. The JKT was administered one task test at a time. The shortest and longest completion times were recorded. At the end of each task test, the contractors discussed each individual task test item with the incumbents. The incumbents were asked to identify any questions, illustrations, or particular words which were difficult to understand. The contractor then identified the correct alternative and asked incumbents for confirmation.

This incumbent pilot test proved valuable for several reasons. One was the fact that because many of the SMEs who assisted in development of the AGE JKT had been in supervisory-type positions for some years, they were not always familiar with current shop language. The incumbents could more readily identify words which would be unfamiliar or confusing to other first-termers. The pilot test also provided an opportunity to obtain an accurate estimate of the amount of time required to complete the test.

Following completion of the incumbent review, the contractor met with a group of four SMEs from Bergstrom AFB who were more senior AGE personnel. The SMEs read through the test, item by item, while the contractor led a discussion on each item. SMEs were asked to consider whether the task was appropriate for the targeted group and if the item content was technically accurate. They were also asked to examine keying, vocabulary, plausible incorrect alternatives, appropriateness and clarity of illustrations, and adequate coverage of tasks. Comments made by the five incumbents were also discussed by the SMEs and the contractor during the review.

This SME review was very valuable because it provided a third opportunity for expert input, thereby enhancing the accuracy and credibility of the test. In addition, it was useful to discuss the test with SMEs from the same unit as the incumbents who provided comments, as these SMEs were in a better position to assure that the incumbent concerns were valid. For example, the incumbents were not familiar with all tasks, such as those dealing with heaters, because Bergstrom is in Texas and heaters are not a commonly maintained piece of equipment. However, some of the SMEs had been stationed at northern bases in the past and, thus, were able to provide meaningful input.

The entire pilot test process took approximately 2 workdays. Eight hours were required for the incumbent review and 8 hours for the SME review.

Item Revision

Approximately 13% of the AGE JKT items were revised based on the information obtained during the pilot test. Items were revised for one of four reasons: (a) to simplify technical terms, (b) to clarify the meaning of questions or options through changes in sentence structure or wording, (c) to include more realistic distractors, and (d) to eliminate poor or implausible distractors. One item was eliminated entirely based on SME recommendations.

Based on the maximum time required to complete tasks during the pilot test, the task tests were divided into two booklets. Booklet A contained 12 tasks consisting of a total of 109 items. The minimum time to complete these items during the pilot test was 41 minutes; the maximum time was 63 minutes. Booklet B contained 14 tasks which were made up of 103 items. The minimum time to complete these items during the pilot test was 42 minutes, and the maximum was 64 minutes. A listing of the tasks and the number of items included for each task can be found in Table 2.

Pretest

The revised AGE JKT was pretested in conjunction with the WTPT pretest at Norton and George AFBs, California. Administration of the two test booklets was counterbalanced at each of the bases in an effort to control for the possible effects of fatigue upon performance on those items toward the end of the test.

Table 2. Number of Items for AGE Task Tests

Task number	Task description	Number of Items
Booklet A		
H260	Clean motor and generator components	5
H264	Isolate engine, motor, or generator malfunctions.	16
I555	Prepare AGE for mobility and training exercises.	8
I488	Remove or install AGE tire, tube and wheel assemblies.	9
I286	Remove or install engine fuel pumps.	7
I255	Change generators and alternators.	11
H215	Perform electrical systems operational checks.	9
H154	Perform an aircraft support generator service inspection.	9
H251	Adjust turbine engine fuel system components.	6
H209	Measure resistance in AGE electrical systems.	9
H179	Perform a gas turbine compressor periodic inspection.	10
I120	Make entries on supply issue and turn-in forms.	10
Booklet B		
I322	Isolate heater system malfunctions.	13
I340	Remove burner control valves on AGE heaters.	6
H284	Remove and replace engine fan belts.	5
I275	Remove or install carburetors.	6
H446	Isolate pneumatic system malfunctions.	5
H503	Research TOs for chassis, enclosures & drive maintenance information.	8
H155	Perform a service inspection on a load bank.	7
H162	Perform a service inspection on a hydraulic test stand.	7
H421	Remove or install hydraulic lines.	8
I181	Perform a hydraulic test stand periodic inspection.	2
I477	Repack wheel bearings.	11
H549	Inspect vehicles for safety of operation.	8
H300	Remove or install fuel lines and fittings.	6
H238	Splice electrical systems wiring.	11

Note. Task numbers correspond to the Occupational Survey tasks from which WTPT items were selected. H indicates a hands-on task; I indicates an interview task.

The examinees were allowed as much time as needed to complete the test. A short rest period was provided after completion of the first test booklet. The shortest test time was approximately 1 hour 45 minutes; the longest time was 4 hours. The majority of examinees completed the test within a 3-hour period. Examinee responses were recorded on an Air Force Type C answer sheet which is machine-scannable and provides five response options (A through E) for each question. The Army used an answer sheet designed specifically for each test. Army answer sheets contained only those response options provided for each question in the test booklet.

The JKT was administered before the WTPT in all cases. The facilities provided for JKT administration were within the requirements described in the administrator's manual. These requirements included adequate lighting, ventilation, and working space. The test was administered to a total of 43 subjects. Subjects were first-term airmen with 6-48 months of on-the-job experience. At George AFB, a total of 27 subjects were tested in six separate testing sessions. At Norton AFB, 16 subjects were tested in two sessions.

A few minor changes were made to items in the JKT as a result of incumbents' comments during the pretest. However, changes were implemented only if substantiated by a specific TO reference. Appendix A provides the characteristics of both the pretest version and the revised final version of the AGE JKT.

Pretest Data Analysis

Pretest data were analyzed in a manner consistent with that used for Army JKTs. An analysis plan ensured the most appropriate information was made available for revising the JKT. For individual JKT items, the number and percentages of subjects selecting each alternative were computed. Also, an item-total score correlation was computed for each item. Total score represented all the items less the subject item in that task test. For the group of items comprising each task test, coefficient alpha was computed.

Based on the results of the pretest it was necessary to reduce the number of AGE JKT items by about 25% because the time allowed for the JKT administration during data collection would be only 2 hours. Item statistics were used to determine which items would be deleted. Guidelines applied by the Army for these purposes were also used in revising the AGE JKT. Any item with a passing rate of less than 10% or greater than 90% was deleted unless inclusion of the item was believed to enhance the "face validity" of the test. The remaining items were examined for low or negative item-total score correlations. These correlations were interpreted with caution due to the heterogeneous nature of the WTPT and the JKT items within each task. Also, some items could not be deleted because they covered an important aspect of the task.

The Army used the Brogden-Ciemans item-total correlation formula (Lord & Novick, 1968) because it corrects for marginal distributions. This correction is necessary when a sample biserial correlation exceeds 1.00. The Brogden-Ciemans formula was not used with these data because the AGE JKT results were not marginally distributed and the resulting correlations did not approach or exceed a value of 1.00. Consequently, a standard point-biserial correlation formula was used. Task tests with low coefficient alphas (indicating low internal consistency) were cautiously examined to determine whether elimination of certain items might improve the reliability level.

Appendix B provides the item analysis results for the pretest version of the AGE JKT. This appendix includes difficulty (percent passing) values and item-total correlations for each of the 212 items in the pretest version of the AGE JKT. Appendix B also shows, by task test, which items were deleted from the pretest version of the test.

After deletion of the weak items, the difficulty levels (percent passing) for individual items ranged from 16% to 95%. Modes for the test items were 70%, 72%, and 81%. Both the median and mean difficulty levels were 58%. These results were similar to those obtained by the Army on their tests; however, the item-total correlations were low in comparison to the Army JKT results. Sixty-seven percent of the item-total correlations for the pretest were positive. However, the average of the item-total correlations in each of the task tests ranged from -.22 to .28.

The deletion of items based on item-total correlations was done with much caution for several reasons. First of all, due to the procedural and mechanical emphasis of the AGE JKT items it is possible that diverse steps within a task may require very different types of knowledge; thus, low correlations would not be unexpected. For example, many tasks included steps which required researching the TOs. A mechanic may be able to accurately perform a mechanical task but be unable to expertly use the TOs associated with that task. In addition, preparation for performance of task steps which includes required safety precautions can be unrelated to the performance of a mechanically oriented task. These negative and low correlations may reflect the heterogeneity of the tests, rather than poor test items.

The AGE JKT task test mean scores ranged from 31% to 86%, and the mean of the task means was 62%. Table 3 provides a frequency distribution of the total test scores. Table 4 provides the mean raw score and standard deviation for each task test. The range of mean percent correct scores for tasks is nearly

identical to those reported by the Army (Campbell et al., 1985). The mean total test score for the 43 subjects was 131.72, with a standard deviation of 10.84. Total test mean percent correct was 62%.

**Table 3. Frequency of AGE (AFS 423X5)
JKT Total Test Scores
(N = 43)**

Total score	Frequency
151 - 155	1
146 - 150	3
141 - 145	5
136 - 140	8
131 - 135	9
126 - 130	6
121 - 125	2
116 - 120	6
110 - 115	3

Note. Maximum possible score = 212.

The coefficient alphas for each task, before deletion of items, varied from -.57 to .55; however, the majority of coefficient alphas were above .20. Table 4 also provides these reliability levels for each task test. After deleting the 53 items listed in Appendix B, coefficient alphas for 14 task tests improved. Table 4 lists the number of items included in each revised task test as well as the estimates of reliability.

Final Revisions

Following the deletion of 53 items from the AGE JKT, the task tests were redistributed between the two booklets. The order of task tests did not change. Booklet A contains 12 tasks with a total of 84 items. Booklet B contains 14 tasks with a total of 75 items.

IV. DEVELOPMENT OF THE AIRCREW LIFE SUPPORT SPECIALIST (AFS 122X0) JOB KNOWLEDGE TEST

The construction of the Aircrew Life Support Specialist JKT took approximately 6 months and was a cooperative effort between three contractor personnel who were simultaneously developing the Personnel Specialist and the Aerospace Ground Equipment Mechanic JKTs.

The general procedures for developing the Aircrew Life Support JKT were: (a) assessment of the WTPT; (b) initial item development, including two SME workshops; (c) pilot test; (d) item revision; (e) pretest; and (f) pretest data analysis.

The Aircrew Walk-Through Performance Test (WTPT)

The WTPT developed for the Aircrew Life Support Specialty (AFS 122X0) includes 23 tasks performed by at least 30% of first-term airmen. Of these 23 tasks, 8 are categorized as Phase I tasks, or those tasks performed across three of the Major Commands, Military Airlift Command (MAC), Strategic Air Command (SAC), and Tactical Air Command (TAC) in the Air Force. The other 15 tasks are part of the Phase II section of the test. Phase II is divided into three different areas, one for MAC, one for SAC, and one for TAC. Five tasks were selected for each of the Phase II areas.

**Table 4. Task Test Analysis for Pretest and Revised Final
Version: Aerospace Ground Equipment Mechanic (AFS 423X5)
Job Knowledge Test
(N = 43)**

Task test	Pretest version				Revised version	
	Number of items	Coefficient alpha	Mean score	SD	Number of items	Coefficient alpha
<u>Booklet A</u>						
260	5	.24	3.33	.92	4	-.16
264	16	.37	11.77	2.01	13	.45
555	8	-.37	5.84	.95	5	-.43
488	9	.22	6.09	1.41	6	.30
286	7	-.08	3.56	1.10	6	-.01
255	11	.10	3.39	1.50	8	.12
215	9	.34	6.51	1.56	8	.41
154	9	.55	7.44	1.55	6	.54
251	6	.18	2.79	1.26	5	.26
209	9	.42	5.70	1.60	7	.60
179	10	-.02	4.95	1.43	7	.12
120	10	.41	5.02	1.75	9	.37
<u>Booklet B</u>						
322	13	.48	6.23	2.24	11	.42
340	6	.35	2.37	1.38	6	.35
284	5	.24	3.60	1.09	4	.32
275	6	.43	3.77	1.15	4	.37
446	5	.01	1.98	1.08	4	.18
503	8	.15	4.07	1.43	7	.28
155	7	.25	5.39	.98	4	.30
162	7	.25	6.02	.91	4	.15
421	8	.23	5.07	1.49	5	.28
181	2	-.57	1.21	.60	2	-.57
477	11	-.46	8.70	1.01	6	-.55
549	8	.31	6.65	.95	4	.29
300	6	.20	3.63	1.20	4	.15
238	11	.35	6.63	1.89	10	.34

As with the AGE WTPT, the Aircrew WTPT includes the same two techniques for extracting job performance information. The first technique is the hands-on performance test in which airmen are asked to perform tasks as they perform them on the job, using any aids typically used such as checklists and TOs. The second technique used is the interview method in which airmen must describe how a task is performed. Of the 23 tasks, 6 were selected as overlap tasks. In these instances, the interview and hands-on tasks are so similar that JKT items were written for the hands-on version only. No tasks were selected for interview testing only due to the nature of the tasks selected for the Aircrew Life Support Specialty. Table 5 lists the tasks included in the WTPT. The maximum time to administer the Aircrew Life Support WTPT is approximately 6 hours 15 minutes.

**Table 5. Aircrew Life Support Specialist (AFS 122X0)
Walk-Through Performance Test Tasks**

Task numbers	Task descriptions
<u>Phase I Tasks</u>	
H199	Make entries on AFTO Form 152.
H295 ^a	Perform final helmet fit using a custom liner.
H315 ^a	Replace the nape strap and pad.
H320 ^a	Remove or replace headsets in helmets.
H330 ^a	Size and fit oxygen masks.
H380	Remove and install filter elements in CRU-80/P.
H383 ^a	Perform mask exchange in the vapor hazard area.
H389 ^a	Fit or adjust parachute harnesses.
<u>Phase II MAC Tasks</u>	
H210	Make entries on AFTO Form 392.
H224	Make entries on DD Form 1574.
H346	Perform oxygen mask connector periodic inspections.
H379	Inspect the MBU-13/P CBO Mask.
H382	In-process an individual wearing the aircrew ensemble through the liquid hazard area.
<u>Phase II SAC Tasks</u>	
H303	Perform 30-day inspection on HGU-55/P helmet.
H310	Perform the nuclear flash blindness goggle kit 180-day inspections.
H349	Perform oxygen mask periodic inspections.
H359	Disassemble, assemble, purge, and refill emergency oxygen cylinders.
H398	Perform 30-day routine parachute inspections.
<u>Phase II TAC Tasks</u>	
H278	Perform the anti-G suit periodic inspection.
H303 ^b	Perform 30-day inspection on a HGU-55/P helmet.
H311	Pour a helmet liner mold.
H349 ^b	Perform oxygen mask periodic inspections.
H483	Repack ACES II survival kits.

Note. Task numbers correspond to the Occupational Survey tasks from which WTPT items were selected; H indicates that the task is a hands-on task in the WTPT; I indicates that the task is an interview task in the WTPT.

^aIndicates those hands-on tasks for which an interview task was also developed.

^bThese tasks are identical to those written for Phase II SAC.

JKT Item Development

The development of the JKT began with six SMEs attending a WTPT workshop conducted at Brooks AFB. Participants jointly identified the key elements of WTPT tasks. A key element was defined as any task step which would have serious repercussions if not performed or if performed incorrectly. After all key elements were identified, the SMEs suggested items that could be written for several of the tasks.

The test developer then used Army item construction guidelines and item writing rules to construct and review new items for all the tasks. One of the main guidelines used when developing items was concentration on the application of knowledge required to perform a task. Therefore, the emphasis was placed on developing performance items. These items required the examinee to actually perform an activity and then identify the obtained answer from item alternatives. However, due to the nature of the tasks on the WTPT, most items had to be performance-related. These items questioned examinees on how something should be performed.

Item writing rules consistent with Army methodology were followed. First, items were written in a multiple-choice format with two to five alternatives per item. Only one of the alternatives could be correct and all distractors had to be plausible responses. Second, most item stems were approximately two lines in length and were worded such that they could be answered without reference to the choices. Third, illustrations were used whenever appropriate in order to convey information in a more effective format and better simulate actual job performance. Fourth, unnecessary technical terms were avoided. Emphasis was placed on using terms normally used on the job. Inter-item cueing was avoided through careful examination.

After construction of first draft items, an internal review took place. Two persons employed by the contractor who were not directly involved in writing the Aircrew JKT per se reviewed each test item. These two people included another test developer and the principal investigator. The review included: (a) an examination for consistency with item writing guidelines, (b) inspection for proper spelling and grammar, (c) consideration of whether WTPT content had been adequately covered, and (d) minimization of individual writing style.

Following the internal review and rewrite, the items were reviewed by another group of six SMEs in a 3-day workshop. The first step in the review process was to once again identify key elements. The SMEs then reviewed the test, task test by task test, item by item. SMEs were told to make sure there was an item for each key element. The items were then reviewed for technical content, understandability, readability, plausible distractors, and correct responses. After all the items were reviewed, the SMEs matched each item to a step or steps from the WTPT. Each step for which an item was not written was then examined to see if an item could be written, in order to ensure complete coverage of the task.

Following this workshop, a second internal review was conducted, particularly targeting newly written items. The items were grouped according to the WTPT task being covered. The task "tests" within each Phase section were ordered from least to most difficult based on difficulty information gathered during WTPT research and in accordance with standard test writing principles. Items within each task test were arranged in the order in which they would occur when actually performed on the job. The test prepared for pilot-testing contained 26 items in 23 task tests.

Pilot Test

This step in the item development process did not serve as a standard or traditional pilot test, but rather as a final review and validation of the items. The pilot test for the Aircrew Life Support Specialty JKT was conducted at Randolph AFB, an Air Training Command (ATC) base in San Antonio, Texas. Four first-term airmen and four SMEs participated in the pilot test.

The first part of the pilot test included only first-term airmen and was labeled incumbent pilot test. The task tests were administered to the first-termers one task test at a time. The performance times required by the first incumbent and the last incumbent to finish each task test were recorded. Following completion of the task test, the test developer discussed each individual task test item with the incumbents. The incumbents were asked what they had chosen as the correct answer. After being told the correct keyed response, they were asked if they agreed that it was the only correct response. Any discrepancies were noted by the test developer. The incumbents were also asked if they understood the questions, the illustrations, and all the words in each question. Feedback was given to the test developer on how questions could be better worded. This entire process was repeated for each task test.

The second stage of pilot-testing involved four SMEs from Randolph AFB who were more senior Aircrew Life Support personnel. This review process was different from that used with the incumbents. The main purpose of using the SMEs during the pilot test was to enhance the accuracy and credibility of the test. The SMEs went through the test item by item. They were told the correct keyed response and were asked if they agreed with the answer. The test developer then led a discussion of each item to ascertain whether: (a) the task was appropriate for the targeted group, (b) the content was technically accurate, (c) alternatives were likely but incorrect, (d) illustrations were appropriate and clear, and (e) there was adequate task coverage.

One difficulty with the pilot test was that the incumbents were not familiar with some of the tasks that were Command specific to MAC, SAC, and TAC. Due to a lack of experience, the incumbents could not provide informed criticism of some specific Phase II tasks. The SMEs, however, were able to give adequate feedback on all tasks because they had been stationed at a variety of bases under all three Commands.

Item Revision

As a result of pilot-testing, 18% of the items were revised based on incumbent and SME criticisms and suggestions. Revisions typically were made for one of four reasons: (a) to simplify technical terms, (b) to clarify the meaning of questions or options through changes in sentence structure or wording, (c) to include more realistic distractors, and (d) to eliminate poor or implausible distractors. Finally, some new items were added to the tests. No items were deleted.

Tasks were not assigned to booklets based on time required to complete the tasks during pilot test as was done with the AGE JKT. Rather, tasks were assigned to booklets according to Phase. All Phase I tasks were included in one booklet. Each set of Phase II tasks was assigned to a separate booklet. Individuals taking the test would take the Phase I booklet and one Phase II booklet, according to their base of assignment. A list of the tasks included in each booklet and the number of items included for each task can be found in Table 6.

Pretest

The pretest was conducted in conjunction with the WTPT pretest at three different bases: (a) Loring AFB, Maine (SAC), (b) Charleston AFB, South Carolina (MAC), and (c) Shaw AFB, South Carolina (TAC). A total of 36 first-term airmen were administered the Phase I Job Knowledge Test; of those 36 airmen, 15 took the SAC Phase II in four testing sessions, 14 took MAC Phase II in seven sessions, and 7 took TAC Phase II in two sessions.

The examinees were allowed as much time as needed to complete the test booklet. A short break was provided after completion of the first test booklet. The mean testing time for the SAC examinees was 84 minutes; for MAC examinees it was 69 minutes; and for TAC examinees, 68 minutes. All of these times were well within the 2-hour period which would be provided for JKT administration during actual data collection.

Order of administration for the Phase I and appropriate Phase II booklet was counterbalanced. A break of 15 minutes was scheduled between booklets in an effort to prevent fatigue. The incumbents responded to the questions on an Air Force Type C answer sheet which is machine-scannable and provides five response options (A through E) for each question.

Table 6. Number of Items for Aircrew Life Support Specialist Task Tests

Task number	Task description	Number of items
<u>Phase I Booklet</u>		
H199	Make entries on AFTO Form 152.	6
H295	Perform final helmet fit using a custom liner.	10
H315	Replace the nape strap and pad.	6
H320	Remove or replace headsets in helmets.	5
H330	Size and fit oxygen masks.	12
H380	Remove and install filter elements in CRU-80/P.	19
H383	Perform mask exchange in the vapor hazard area.	8
H389	Fit or adjust parachute harnesses.	4
<u>Phase II MAC Booklet</u>		
H210	Make entries on AFTO Form 392.	4
H224	Make entries on DD Form 1574.	7
H346	Perform oxygen mask connector periodic inspections.	11
H379	Inspect the MBU-13/P CBO Mask.	15
H382	In-process an individual wearing the aircrew ensemble through the liquid hazard area.	15
<u>Phase II SAC Booklet</u>		
H303	Perform 30-day inspection on HGU-55/P helmet.	7
H310	Perform the nuclear flash blindness goggle kit 180-day inspections.	14
H349	Perform oxygen mask periodic inspections.	20
H359	Disassemble, assemble, purge, and refill emergency oxygen cylinders.	18
H398	Perform 30-day routine parachute inspections.	14
<u>Phase II TAC Booklet</u>		
H278	Perform the anti-G suit periodic inspection.	9
H303	Perform 30-day inspection on a HGU-55/P helmet.	7
H311	Pour a helmet liner mold.	12
H349	Perform oxygen mask periodic inspections.	20
H483	Repack ACES II survival kits.	20

Appendix C provides a detailed description of the Aircrew Life Support JKT pretest. The Appendix also describes the final version after changes were made as a result of pretest.

Pretest Data Analysis

Analyses were run on both individual item responses and task tests. For each item written, the number and percentage who selected each alternative were calculated. Those items that were too easy (above 95% passing) or too difficult (below 10% passing) were examined closely for obvious cueing or keying errors. If these types of errors were found, the item was revised. If neither of these errors could explain the high or low pass rate, the item was eliminated. A few items in the test were duplicated due to the overlap of tasks in the Phase II tests. These items (those repeated on more than one task test) were eliminated from the test as well.

An item-total score correlation was also computed in order to assess the homogeneity of items within a task test. Low or negative correlations were looked at carefully. For the most part, these correlations were used only when making borderline judgments on whether to keep an item. Although statistics were computed for Phase I and II items, only those calculated for Phase I were used to make decisions concerning the elimination or reduction of items. The small sample sizes for the Phase II tests prohibited use of the analyses in test revision.

A total of five items were deleted from Phase I of the Aircrew Life Support JKT. Two of the items were deleted because they appeared in two task tests. The other three were deleted because of low item difficulty levels. The Phase II tests for MAC, SAC, and TAC were not revised. Appendix D shows which items were deleted from Phase I of the test.

Following deletion of the five items, item statistics were recalculated. For Phase I, the difficulty level, or percent passing, ranged from 14% to 94%, with a mean percent passing of 57% and a median of 60%. For Phase II MAC items, percent passing ranged from 0% to 100%, with a mean of 51% and a median of 50%. The range of difficulty levels for Phase II SAC items was 7% to 100%. The mean was 64%; the median, 73%. Phase II TAC items had difficulty levels ranging from 0% to 100%. The mean was 70% and the median was 71%.

Appendix D gives the pretest item analysis results for Phase I, including the number and percentages selecting each alternative for each item and the item-total correlations. Appendix E presents the pretest item analysis for Phase II, including the number selecting each response and the item-total correlations. Because of the small sample sizes for the Phase II portions of the test, percentages are not reported.

The item-total correlations for Phase I items ranged from $-.36$ to $.55$, with positive correlations for 53 (76%) items. The Phase II MAC test had 36 (71%) items with positive item-total correlations; correlations ranged from $-.60$ to $.74$. The Phase II SAC test had 48 (67%) items with positive item-total correlations; correlations ranged from $-.42$ to $.89$. Forty-eight percent of the Phase II TAC item-total correlations were positive, and correlations ranged from $-.75$ to $.94$. The Aircrew Life Support tasks were procedural and heterogeneous in nature, requiring that negative correlations be examined with caution. Diverse knowledge can be required in order to successfully complete the steps of some tasks. Also affecting the correlations is the test length; task tests ranged from 4 to 20 questions.

The average item-total correlations for task tests ranged from $.01$ to $.30$ for Phase I, with an overall average of $.16$. The average item-total correlations for the MAC Phase II test ranged from $-.01$ to $.40$; for the SAC Phase II test the range was $.00$ to $.34$; and for the TAC Phase II test, $-.31$ to $.31$.

Task test analyses were also performed. Included were task test means, standard deviations, and coefficient alphas. Following elimination of items, coefficient alphas were recomputed. As with the item analyses, all of the task test statistics were run for both Phase I and Phase II task tests.

Table 7 presents means, standard deviations, and coefficient alphas for Phase I pretest task tests and coefficient alphas for the revised Phase I test. Table 8 presents the means, standard deviations, and coefficient alphas for the Phase II portions of the test.

**Table 7. Task Test Analysis for Pretest and Revised Final Versions:
Aircrew Life Support Specialist (AFS 122X0)
Job Knowledge Test: Phase I (N = 36)**

Task test	Pretest version				Revised version	
	Number of items	Coefficient alpha	Mean score	SD	Number of items	Coefficient alpha
389	4	-.03	2.94	.71	3	-.03
199	6	.08	4.61	1.02	6	.08
315	6	.54	4.11	1.41	5	.63
320	5	.06	2.50	1.08	5	.06
383	8	.5	5.22	1.73	8	.51
330	12	.52	8.94	2.03	12	.52
380	19	.69	10.42	3.41	18	.68
295	10	.28	5.36	1.62	8	.25

**Table 8. Task Test Analysis for Pretest and Revised Final Versions:
Aircrew Life Support Specialist (AFS 122X0)
Job Knowledge Test: Phase II (N = 14)**

Task test number	Number of items	Coefficient alpha	Mean score	SD
Phase II MAC (N = 14)				
346	11	.75	5.14	2.80
379	15	-.10	7.50	1.65
382	15	.65	6.29	2.43
210	4	-.02	1.07	.83
224	7	.47	5.14	1.35
Phase II SAC (N = 15)				
303	7	.56	4.13	1.50
349	20	.60	14.60	2.44
398	14	.12	9.87	1.73
359	18	.75	10.80	3.65
310	14	-.09	7.27	1.53
Phase II TAC (N = 7)				
303	7	-.19	4.29	.76
349	20	-.48	15.14	1.21
311	12	.63	7.14	2.41
278	9	.20	8.14	.90
483	20	.75	12.86	3.80

Table 9 presents a frequency distribution of raw scores for Phase I of the test. The mean raw score is 44, with a standard deviation of 3.99. Phase I test mean percent correct was 63%.

Table 9. Frequency of Aircrew Life Support Specialist (AFS 122X0) Phase I JKT Total Test Scores

Total score	Frequency
51 - 55	7
46 - 50	13
41 - 45	3
36 - 40	6
31 - 35	1
25 - 30	6

Note. Maximum possible score = 70.

The maximum possible scores for the three separate Aircrew Life Support tests were as follows: (a) MAC - 122, (b) SAC - 143, and (c) TAC - 138. The mean total score for MAC was 62, with a range from 45 to 84. The mean percent correct was 51%. The mean total score for SAC was 97, with a range of 73 to 122. The mean percent correct was 68%. For TAC, the scores ranged from 83 to 105, with a mean of 92.86 and a mean percent correct rate of 67%. The frequency distributions for Phase II test scores are shown in Table 10.

Table 10. Frequency, Mean, and Standard Deviation of Aircrew Life Support Specialist (AFS 122X0) JKT Total Test Scores

	MAC (N = 14)	SAC (N = 15)	TAC (N = 7)
Max. possible score	122	143	138
Total score	Frequency		
116 - 125		1	
106 - 115		3	
96 - 105		5	2
86 - 95		3	3
76 - 85	2	1	2
66 - 75	3	2	
56 - 65	3		
45 - 55	6		
Mean	62	97	93
Standard Deviation	12.10	13.49	7.30
Mode	51	105, 107	83, 93
Median	61	98	93
Mode % Correct	42%	73%, 75%	60%, 67%
Mean % Correct	51%	68%	67%

V. DEVELOPMENT OF THE PERSONNEL SPECIALIST (AFS 732X0) JOB KNOWLEDGE TEST

The construction of the Personnel JKT took approximately 5 months. The Personnel JKT development began about 1 month after the start of AGE and Aircrew Life Support JKTs development. It was a cooperative effort between three contractor personnel who were simultaneously developing the AGE and Aircrew Life Support JKTs.

The general procedures for developing the Personnel JKT were: (a) assessment of the WTPT; (b) initial item development, including two SME workshops; (c) pilot test; (d) item revision; (e) pretest; and (f) pretest data analysis.

The Personnel Walk-Through Performance Test (WTPT)

The Personnel Specialist WTPT, which uses both hands-on measures of performance and an incumbent interview technique, consists of a total of 71 tasks. Like the Aircrew Life Support WTPT, the Personnel WTPT is divided into two Phases. Phase I of the test includes a set of 8 tasks which are representative of tasks performed by everyone in the job specialty. There are five Phase II sections. Each contains from 11 to 13 tasks and represents a different functional area within the specialty. These five functional areas are: (a) Manning Control, (b) Classification and Training, (c) Records, (d) Outbound Assignments, and (e) Separations. Out of the 71 tasks, 37 are hands-on tasks and 34 are interview tasks. Thirteen of the 71 tasks are overlap tasks; i.e., both a hands-on and an interview test were written for the task. Job knowledge task tests were written for only the hands-on version of the overlap tasks. Therefore, JKT task tests were written for a total of 58 tasks. Table 11 provides a list of the tasks included in the Personnel WTPT, dividing them into their appropriate sections. The table also identifies each task as either a hands-on, interview or overlap task. Each examinee is administered the Phase I test and the one Phase II portion that corresponds with his/her work center. The maximum time required to administer the entire test depends on the work center for which an examinee is being tested. The maximum times required to complete the WTPT for each work center are as follows: (a) Manning Control, 2 hours 34 minutes; (b) Classification and Training, 2 hours 5 minutes; (c) Records, 2 hours 20 minutes; (d) Outbound Assignments, 2 hours 39 minutes; and (e) Separations, 1 hour 53 minutes.

JKT Item Development

Item development began with a workshop attended by six Personnel SMEs. As in other SME workshops, the SMEs reviewed the WTPT tasks as a group, identifying those steps judged to be key elements. SMEs also identified plausible incorrect methods for performing those steps identified as key elements. Following this workshop, items were drafted for the key elements in each task. The developmental procedure took into account all information learned during the workshop, as well as information obtained from WTPT task analysis, AFRs, and CDCs. After items were developed to represent the key elements, additional items were constructed to achieve adequate coverage of each task.

Again, items were developed to concentrate on the application of the knowledge needed to perform a task. In addition, rules consistent with Army JKT development methodology were followed during item construction. All items were written in a multiple-choice format with two to five alternatives per item. The number of alternatives was directly related to the number of plausible incorrect responses. Alternatives were written such that only one could be considered correct. Item stems were usually limited to two lines and were worded, whenever possible, such that they could be answered without reference to the choices. Illustrations were used as much as possible to convey information in a more efficient and effective format and to better simulate job performance. An emphasis was placed upon using terms or acronyms which are commonly used on the job in an effort to ensure that an examinee's performance was being tested and not knowledge of materials and terms, or reading level. Items were also carefully examined to eliminate any inter-item cueing.

Table 11. Personnel Specialist (AFS 732X0) WTPT Tasks

Task number	Task description
<u>Phase I</u>	
I733 ^a	Release information from a member's file in accordance with the Privacy Act.
I121	Dispose of PDS products.
H131	Maintain charge out record forms (AF Form 614).
H35	Draft a message.
H896	Compute service data.
H140	Open and close CRT.
H719 ^a	File documents in personnel records.
H116	Construct immediate inquiries.
<u>Phase II Manning Control</u>	
I447	Distribute allocation briefs.
I437	Report discrepancies after auditing incoming records for receipt of items reflected on the Records Transmittal/Request.
I475	Prepare incoming PCS processing folders.
H472	Perform duty change actions.
H466 ^a	Update overseas data.
I441	Conduct in-processing briefings.
H440	Check personnel reporting dates for overdue personnel.
H436	Assemble manning statistics.
H476 ^a	Generate an IDA RIP.
HA1 ^a	Assign personnel to duty positions.
<u>Phase II Classification and Training</u>	
I335	Process retraining declination statements.
I324	Prepare applications for retraining.
I343	Ensure personnel maintain qualifications for attendance at NCO Leadership School.
H303	Establish officer upgrade suspense actions for award of fully qualified AFSCs.
H157	Prepare request and authorization for temporary duty forms.
I296	Conduct classification interviews during in-processing.
E145 ^a	Perform temporary duty relocation actions.
HADD	Process formal training allocation cancellations.
H293 ^a	Process awarded officer AFSCs.
H334	Process AFSC withdrawal actions.
<u>Phase II Records</u>	
I734	Remove obsolete documents from records.
I728 ^a	Prepare statements of service.
H739	Verify data provided on VA Forms (for educational purposes).
I722	Perform periodic records inventories.
I711	Audit records reviews listings (RRL).
I710	Assemble senior NCO promotion selection folders.
H720	File personnel records folders.

Table 11. (Concluded)

Task number	Task description
I735	Review official photographs for file.
H713	Conduct records reviews.
H718 ^a	Establish computer records for prior service members.

Phase II Outbound Assignments

H406	Process applications for concurrent travel.
I370	Maintain clearance record files on personnel who have departed PCS.
H357	Conduct initial assignment briefings.
H433	Verify completion of assignment relocation processing.
H415 ^a	Complete PCS or TDY declination statements (AF Form 964).
H398	Prepare request and authorize PCS-military forms (AF Form 899).
H389	Prepare nomination documents for personnel being assigned to SCI positions (AF Form 14).
H388	Prepare medical and educational clearance for dependent overseas travel forms (AF Form 1466).
H396 ^a	Prepare request and authorization for change of administrative order forms (AF Form 973).
H380 ^a	Prepare assignment reclaims.

Phase II Separations

I861	Maintain separations preparation project folders.
I839	Collect I.D. cards from separatees or retirees.
H889	Process hardship discharge requests.
H884	Prepare separation orders.
H876 ^a	Prepare pregnancy discharge requests.
I852	Evaluate PALACE CHASE separation requests.
I840	Conduct final separation briefings or interviews.
H835 ^a	Advise first-term four-year enlistee on military service dates.
H878	Prepare report of separation from active duty forms (DD Form 214).
H874 ^a	Prepare PALACE CHASE "Category A" unit selected reserve service contract forms (AF Form 3028).

Note. Task numbers correspond to the Occupational Survey tasks from which WTPT items were selected; H indicates that the task is a hands-on task in the WTPT.; I indicates that the task is an interview task in the WTPT.

^aIndicates overlap tasks.

After completion of this initial item development stage, the items were reviewed during a second Personnel SME workshop attended by five SMEs. In order to review all items in the 3 days provided for the SME workshop, the SMEs were divided into two different groups each day. One work center was reviewed by each group, so that two work centers could be covered each day. Phase I was considered and reviewed as a work center section. During a day's review, the SMEs were guided through a discussion of the written items. The SMEs were instructed to consider item content for technical accuracy and correct wording. In addition, they were asked to confirm correct answers, determine whether the distractors were plausible, and generate additional distractors when needed. SMEs examined the test's illustrations for accuracy and also

considered whether or not any items could be strengthened by using illustrations. Finally, SMEs determined whether each task was sufficiently covered by the items written or if additional items needed to be developed.

Following this second workshop, a first draft of the test was composed, taking into account all the information obtained from the two groups of SMEs. Each item written was reviewed by two persons employed by the contractor. One of the two reviewers was the principal investigator on the project; the other was the research assistant who was simultaneously writing JKT items for the other work centers in the Personnel test. The reviewers looked for consistency with item writing guidelines and inspected the items for proper spelling, grammar, and readability. The reviewers tried to determine whether WTPT tasks had been adequately covered by paper-and-pencil items. The review also involved editing to minimize the personal writing style of the principal test developer.

Following the internal review of the Personnel JKT, an informal review of items was conducted at Brooks AFB, TX. Each SME was responsible for reviewing a section of the Personnel JKT. One JKT developer led the discussion. The SMEs were asked to identify the correct responses and any problem areas.

Once constructed and reviewed, the items were grouped according to the WTPT task being covered. The "task tests" were then grouped into booklets according to work center. The Personnel JKT tasks often involved the use of forms, records, or computer printouts such as Records of Individual Personnel (RIPs). These documents were compiled in a separate supplementary book for each work center so that the JKT items and reference pages could be more conveniently maneuvered by the examinees during test administration. Therefore, the Personnel JKT consists of a total of eleven booklets. One booklet contains Phase I tasks, five booklets contain the Phase II tasks for the five work centers, and five booklets contain the supplementary materials for the five work centers.

The tasks within each test booklet were ordered from least difficult to most difficult based on difficulty information gathered during WTPT research. The test contained a total of 301 items in 58 task tests. Table 12 lists the tasks and number of items comprising the Phase I booklet and each work center booklet of the Personnel JKT.

Pilot Test

This step in the item development process did not serve as a standard or traditional pilot test of items, but rather as a final review and validation of the items. The Personnel JKT was pilot-tested at Loring AFB, Maine. Six incumbents (first-term airmen) and eight SMEs attended the same session. At least one SME and one incumbent completed each phase of the test. Both the incumbents and the SMEs were instructed to record any problems or to point them out to one of the two research assistants conducting the pilot test. They were also provided with an answer key after completion of the test to verify their answers.

The incumbents and SMEs identified questions, illustrations, or particular words they found difficult to understand. They also considered whether each task had been adequately covered.

Though the procedure followed for the Personnel pilot test differed somewhat from the AGE and Aircrew pilot test procedures, valuable information was obtained. The pilot test provided some information concerning procedural changes in some of the task tests which had not yet been implemented by all bases. It also provided the first incumbent review of the Personnel test. This incumbent review was valuable because it provided input on the technical language used by the targeted population. The incumbent review also provided an opportunity for estimating the time required to complete the test. Though performance times were not recorded for individuals, everyone finished well within the 2-hour period which would be allowed during data collection.

**Table 12. Number of Items for Personnel Specialist
Task Tests**

Task number	Task description	Number of items
<u>Phase I</u>		
I733	Release information from a member's file in accordance with the Privacy Act.	5
I121	Dispose of PDS products.	2
H131	Maintain charge out record forms (AF Form 614).	5
H35	Draft a message.	9
H896	Compute service data.	3
H140	Open and close CRT.	3
H719	File documents in personnel records.	10
H116	Construct immediate inquiries.	10
<u>Phase II Manning Control</u>		
I447	Distribute allocation briefs.	2
I437	Report discrepancies after auditing incoming records for receipt of items reflected on the Records Transmittal/Request.	8
I475	Prepare incoming PCS processing folders.	2
H472	Perform duty change actions.	4
H466	Update overseas data.	3
I441	Conduct in-processing briefings.	7
H440	Check personnel reporting dates for overdue personnel.	5
H436	Assemble manning statistics.	6
H476	Generate an IDA RIP.	5
HA1	Assign personnel to duty positions.	5
<u>Phase II Classification and Training</u>		
I335	Process retraining declination statements.	4
I324	Prepare applications for retraining.	5
I343	Ensure personnel maintain qualifications for attendance at NCO Leadership School.	6
H303	Establish officer upgrade suspense actions for award of fully qualified AFSCs.	4
H157	Prepare request and authorization for temporary duty forms.	6
I296	Conduct classification interviews during in-processing.	5
H145	Perform temporary duty relocation actions.	7
HADD	Process formal training allocation cancellations.	5
H293	Process awarded officer AFSCs.	6
H334	Process AFSC withdrawal actions.	5
<u>Phase II Records</u>		
I734	Remove obsolete documents from records.	5
I728	Prepare statements of service.	3
H739	Verify data provided on VA Forms (for educational purposes).	3
I722	Perform periodic records inventories.	2
I711	Audit records reviews listings (RRL).	5

Table 12. (Concluded)

Task number	Task description	Number of items
I710	Assemble senior NCO promotion selection folders.	8
H720	File personnel records folders.	4
I735	Review official photographs for file.	3
H713	Conduct records reviews.	5
H718	Establish computer records for prior service members.	3

Phase II Outbound Assignments

H406	Process applications for concurrent travel.	3
I370	Maintain clearance record files on personnel who have departed PCS.	3
H357	Conduct initial assignment briefings.	5
H433	Verify completion of assignment relocation processing.	3
H415	Complete PCS or TDY declination statements (AF Form 964).	6
H398	Prepare request and authorize PCS-military forms (AF Form 899).	14
H389	Prepare nomination documents for personnel being assigned to SCI positions (AF Form 14).	11
H388	Prepare medical and educational clearance for dependent overseas travel forms (AF Form 1466).	4
H396	Prepare request and authorization for change of administrative order forms (AF Form 973).	5
H380	Prepare assignment reclaims.	4

Phase II Separations

I861	Maintain separations preparation project folders.	4
I839	Collect I.D. cards from separatees or retirees.	3
H889	Process hardship discharge requests.	2
H884	Prepare separation orders.	8
H876	Prepare pregnancy discharge requests.	5
I852	Evaluate PALACE CHASE separation requests.	4
I840	Conduct final separation briefings or interviews.	4
H835	Advise first-term four-year enlistee on military service dates.	7
H878	Prepare report of separation from active duty forms (DD Form 214).	13
H874	Prepare PALACE CHASE "Category A" unit selected reserve service contract forms (AF Form 3028).	5

Item Revision

Approximately 17% of the Personnel JKT items were revised based upon the information obtained during pilot test. Items were revised for one of four reasons: (a) to simplify technical terms, (b) to clarify the meaning of questions or options through changes in sentence structure or wording, (c) to include more realistic distractors, and (d) to eliminate poor or implausible distractors.

The revised JKT (to be used in pretest) still consisted of 58 task tests with a total of 301 items. The number of items in each task test or phase did not change. Appendix F provides a detailed description of the Personnel JKT characteristics.

Pretest

The Personnel JKT was pretested with 25 airmen at two Air Force bases. Thirteen subjects completed the test at Charleston AFB, South Carolina, and 12 individuals took the test at Pease AFB, New Hampshire. The Personnel WTPT had been administered to the same subjects several weeks earlier.

Two testing sessions were held at each base; no break was required during these sessions due to the shorter length of the Personnel JKT. Administration of the Phase I and Phase II booklets was counterbalanced, and each airman completed Phase I and the Phase II booklet corresponding to his/her current work center. One individual who had experience in two different work centers completed Phase II booklets for both. Phase II booklets were completed by three airmen from the Manning Control work center, three from Classification and Training, seven from Records, six from Outbound Assignments, and seven from Separations. Subjects responded to questions on the Air Force Type C machine-scannable answer sheet.

No time limits were imposed during administration. The entire Personnel JKT required approximately 40 to 80 minutes to complete. Phase I took approximately 15 to 30 minutes. Phase II booklets were completed in 20 to 40 minutes.

Pretest Data Analysis

Personnel pretest data were analyzed according to a plan which was consistent with the Army's data analysis plan. It was not necessary to reduce the number of Personnel JKT items because the test was already within the 2-hour period provided for test administration during data collection.

Results from Phase I tasks and items were analyzed, but due to the small sample size, results were interpreted cautiously. For individual items, the number and percentage of subjects selecting each alternative were computed. In addition, an item-total correlation was computed for each item within a task test. Total score represented all the items less the subject item in that task test. For the group of items comprising each task test, coefficient alpha was computed. Results from the Phase II test administration were not used in test revision due to the extremely small sample sizes.

The item analysis results for the pretest version of the Phase I Personnel test are presented in Appendix G. This appendix includes difficulty (percent passing) values and item-total correlations for each of the 47 items in the pretest version of the Phase I Personnel JKT. The difficulty levels (percent passing) for individual items ranged from 20% to 100%. The median difficulty level was 84% and the mean was 75%. These difficulty levels were higher than those obtained by the Army on the Administrative Specialist JKT. This MOS involved job tasks similar to those in the Personnel specialty. These results suggest that the Personnel test is easier than the Army test; however, such a decision should be made with caution. The Personnel test results on the 47 items in Phase I were based on a sample of only 25 subjects. The Army's Administrative Specialist JKT had 170 items and analysis was performed using a sample of 129 subjects (Campbell et al., 1985).

The Personnel JKT task test mean scores ranged from 51% to 89%, and the mean of the task means was 67%. Table 13 provides the mean raw score and standard deviation for each Phase I task test. The mean total test score for the 25 subjects was 35.96, with a standard deviation of 4.54. Total test mean percent correct was 75%. Table 14 provides a frequency distribution of the Phase I test scores.

The coefficient alphas for each task in Phase I varied from .34 to .90. This information should also be interpreted cautiously due to the sample sizes. Table 13 provides the reliability levels for each of the task tests.

The Phase II test data were not analyzed. The small sample sizes for each work center prohibited useful interpretation of the results.

**Table 13. Task Test Analysis for Pretest and Revised Final
Versions: Personnel Specialist (AFS 732X0)
Job Knowledge Test: Phase I
(N = 25)**

Task test number	Number of items	Coefficient alpha	Mean score	SD
I733	5	.50	2.54	1.36
I121	2	.34	1.04	.60
H131	5	.72	3.96	1.34
H35	9	.79	7.42	2.00
H896	3	.35	1.54	.99
H140	3	.55	1.77	.99
H719	10	.90	8.85	2.33
H116	10	.67	7.31	2.09

**Table 14. Frequency of Personnel Specialist
(AFS 732X0) JKT Phase I Total Test Scores
(N = 25)**

Total score	Frequency
43 - 45	1
40 - 42	5
37 - 39	7
34 - 36	6
31 - 33	2
28 - 30	4

Note. Maximum possible score = 47.

VI. CONCLUSION

The objective of the present R&D effort was to construct paper-and-pencil job knowledge tests for three Air Force specialties (Aerospace Ground Equipment Mechanic, Aircrew Life Support Specialist, and Personnel Specialist), using procedures developed by the Army. The research was initiated to determine whether measurement technology developed in one Service could be effectively transferred to another Service. The objective was accomplished with only slight deviations from Army methodology. The three tests were developed after careful review of all Army procedures and substantial contact with SMEs from the three Air Force career fields addressed.

With few exceptions, the resulting tests appeared to be similar to those developed by the Army. Two of the Air Force tests had fewer illustrations than the Army tests did; thus, future endeavors might utilize an on-staff artist. Because the amount of time available for test development during this project was constrained, employing an artist from another organization would have been too time-consuming. The type of answer sheet used by the Army was specially designed to correspond with the varying number of response options for Army test items. The tests developed in this study employed standard machine-scannable answer sheets with a standard number of response items, as opposed to the answer sheets developed by the Army specifically for the JKT, which had the number of responses on the answer sheet equal to that of the individual item. Another difference, of a cosmetic nature, was in the numbering of test items. The Army began numbering items in each task test with the number one. The Air Force, because of the answer sheet used, numbered items consecutively from task test to task test.

The final versions of the tests will be administered to much larger samples than those used within this effort. The results from the full-scale data collection effort should be more enlightening than those reported here.

Pretest results suggest, for the most part, that the tests are good measures of the tasks included in the WTPT. A thorough review of the full-scale data collection results will provide a basis for making more definitive observations concerning the accomplishments of this research effort.

REFERENCES

- Augustin, J.W., Bentley, B.A., & Ringenbach, K. (1987a). *Army job performance measures: A description of developmental procedures*. San Antonio, TX: Universal Energy Systems, Inc.
- Augustin, J.W., Bentley, B.A., & Ringenbach, K. (1987b). *Strategy for development of the Job Knowledge Test for Aerospace Ground Equipment (AFS 423X5)*. San Antonio, TX: Universal Energy Systems, Inc.
- Augustin, J.W., Bentley, B.A., & Ringenbach, K. (1987c). *Strategy for development of the Job Knowledge Test for Personnel Specialist (AFS 732X0)*. San Antonio, TX: Universal Energy Systems, Inc.
- Augustin, J.W., Ringenbach, K., & Bentley, B.A. (1987). *Strategy for development of the Job Knowledge Test for Aircrew Life Support (AFS 122X0)*. San Antonio, TX: Universal Energy Systems, Inc.
- Campbell, C.H., Campbell, R.C., Rumsey, M.G., & Edwards, D.C. (1985). *Development and field test of task-based MOS-specific criterion measures* (Tech. Rep. No. 717). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Lord, F., & Novick, M. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.
- Washington, DC: Headquarters Department of the Army. (1982). *Soldier's manual of common tasks skill level 1 (FM21-2)*.
- Washington, DC: Headquarters Department of the Army (1985). *Soldier's manual for 71L, skill levels 1/2, administrative specialist (STP 12-71L12-SM)*.

**APPENDIX A: AEROSPACE GROUND EQUIPMENT PRETEST
AND FINAL TEST CHARACTERISTICS**

Characteristics	Pretest version		Final version	
# of Booklets	2		2	
# of Task Tests	26		26	
# of Items	212		159	
% of Items with Illustrations	15%		18%	
Range of Items per Task Test	2-16		2-13	
Distribution of Number of Items per Task Test	Items	Tasks	Items	Tasks
	2	1	2	1
	5	3	4	8
	6	4	5	3
	7	3	6	5
	8	4	7	3
	9	4	8	2
	10	2	9	1
	11	3	10	1
	13	1	11	1
	16	1	13	1
Average Number Items per Task	8.15		6.12	
Range of Number of Choices per Item	2-5		3-5	
Distribution of Choices	Choices	Items	Choices	Items
	2	1	3	35
	3	50	4	118
	4	154	5	
	5	7		
Printed Pages	83		71	
Time to Finish Test (Range)	2-4 hours		unknown	

Note. No description of the pilot test is reported due to the unstandardized nature of the pilot test.

**APPENDIX B: ITEM ANALYSIS FOR PRETEST VERSION OF
THE AGE (AFS 423X5) JOB KNOWLEDGE TEST (N = 43)**

Item	Response Alternative										r ^c
	A		B		C		D		E		
	N	%	N	%	N	%	N	%	N	%	
Task Test #260											
1	1	2	13	30	29	67 ^b					-.22
2 ^a	2	5	39	91 ^b	1	2					-.12
3	18	42 ^b	14	33	3	7	8	19			.04
4	3	7	3	7	3	7	32	74 ^b	2	5	.03
5	3	7	1	2	14	33	25	58 ^b			-.15
Task Test #264											
6	1	2	1	2	35	81 ^b	6	14			.28
7	1	2	32	74 ^b	1	2	9	21			.12
8 ^a	0	0	38	88 ^b	0	0	0	0	5	12	-.30
9	0	0	30	70 ^b	5	12	8	19			.19
10 ^a	1	2	1	2	41	95 ^b	0	0			.09
11	1	2	37	86 ^b	1	2	4	9			.16
12	2	5	23	54 ^b	8	19	10	23			.22
13 ^a	0	0	2	5	41	95 ^b					-.02
14	6	14	0	0	36	84 ^b	1	2			.11
15	5	12	20	47 ^b	5	12	13	30			.15
16	37	86 ^b	2	5	3	7	1	2			.23
17	13	30	15	35 ^b	1	2	14	33			-.20
18	0	0	0	0	37	86 ^b	6	14			.31
19	1	2	0	0	7	16	35	81 ^b			.28
20	31	72 ^b	4	9	5	12	3	7			.15
21	4	9	18	42	18	42 ^b	3	7			.09
Task Test #555											
22	4	9	24	56 ^b	14	33	1	2			-.09
23 ^a	40	93 ^b	1	2	0	0	2	5			-.03
24 ^a	0	0	0	0	43	100 ^b	0	0			.00
25 ^a	2	5	1	2	40	93 ^b					-.03
26	14	33 ^b	5	12	23	54					-.31
27	1	2	0	0	7	16	35	81 ^b			-.18
28	16	37	21	49 ^b	6	14					-.07
29	5	12	34	79 ^b	3	7	1	2			.03
Task Test #488											
30 ^a	1	2	41	95 ^b	1	2					-.06
31	6	14	0	0	31	72 ^b	6	14			.28
32	0	0	1	2	15	35	27	63 ^b			.22
33	8	19	2	5	25	58 ^b	8	19			-.10
34	4	9	36	84 ^b	1	2	2	5			.29
35	12	28 ^b	13	30	5	12	13	30			.27
36 ^a	24	56	0	0	18	42 ^b	1	2			-.04
37 ^a	39	91 ^b	3	7	1	2					.10
38	0	0	3	7	33	77 ^b	7	16			-.19

APPENDIX B: (Continued)

Item	Response Alternative										r
	A		B		C		D		E		
	N	%	N	%	N	%	N	%	N	%	
Task Test #286											
39 ^a	0	0	0	0	42	98 ^b	1	2			-.20
40	21	49 ^b	1	2	8	19	13	30			-.11
41	34	79	7	16 ^b	2	5					.01
42	1	2	15	35 ^b	0	0	27	63			.10
43	2	5	3	7	35	81 ^b	3	7			.06
44	19	44 ^b	5	12	9	21	10	23			-.01
45	8	19	6	14	14	33 ^b	15	35			-.11
Task Test #255											
46	1	2	19	44 ^b	2	5	21	49			.25
47 ^a	23	54	6	14	13	30 ^b	1	2			.03
48	0	0	12	28 ^b	11	26	20	47			-.01
49	10	23 ^b	4	9	19	44	10	23			.27
50	11	26	31	72 ^b	1	2					.00
51	2	5	15	35	12	28	14	33 ^b			-.20
52	9	21 ^b	16	37	4	9	14	33			.14
53	6	14	11	26	12	28 ^b	14	33			.18
54	5	12	10	23	20	47 ^b	8	19			-.18
55 ^a	5	12	6	14 ^b	7	16	25	58			-.07
Task Test #215											
57	1	2	30	70 ^b	12	28	0	0			.24
58	27	63 ^b	4	9	1	2	11	26			.18
59	5	12	33	77 ^b	5	12	0	0			.26
60 ^a	0	0	3	7	39	91 ^b	1	2			-.18
61	2	5	3	7	2	5	36	84 ^b			.12
62	7	16	22	51 ^b	6	14	8	19			-.12
63	30	70 ^b	3	7	7	16	3	7			.24
64	27	63 ^b	2	5	2	5	12	28			.18
65	4	9	36	84 ^b	1	2	2	5			.26
66 ^a	41	95 ^b	0	0	2	5	0	0			.15
67	33	77 ^b	8	19	2	5					.32
68	1	2	37	86 ^b	0	0	5	12			.12
69 ^a	0	0	2	5	40	93 ^b	1	2			.36
70	29	67 ^b	1	2	11	26	2	5			.37
71	7	16	1	2	34	79 ^b	1	2			.43
72	28	65 ^b	5	12	1	2	9	21			.37
73 ^a	1	2	41	95 ^b	0	0	1	2			.07
74	0	0	3	7	37	86 ^b	3	7			.07

APPENDIX B: (Continued)

Item	Response Alternative										r
	A		B		C		D		E		
	N	%	N	%	N	%	N	%	N	%	
Task Test #251											
75	5	12	6	14	21	49 ^b	11	26			.26
76	11	26	13	30 ^b	7	16	12	28			.03
77 ^a	9	21 ^b	21	49	8	19	5	12			-.06
78	15	35	20	47 ^b	6	14	2	5			-.06
79	1	2	8	19	4	9	30	70 ^b			.10
80	27	63 ^b	11	26	4	9					.21
Task Test #209											
81	17	40	2	5	21	49 ^b	3	7			.26
82	4	9	36	84 ^b	0	0	3	7			.22
83 ^a	38	88 ^b	3	7	2	5	0	0			-.22
84	8	19	4	9	31	72 ^b	0	0			.42
85	5	12	22	51 ^b	16	37	0	0			.35
86	35	81 ^b	3	7	5	12					.34
87	4	9	3	7	1	2	34	79 ^b			.29
88 ^a	13	30	17	40	3	7 ^b	10	23			-.38
89	6	14	12	28	25	58 ^b					.10
Task Test #179											
90	10	23	30	70 ^b	1	2	2	5			.13
91	6	14	16	37	2	5	19	44 ^b			-.03
92 ^a	0	0	0	0	42	98 ^b	1	2			.22
93	7	16	11	26	13	30 ^b	12	28			-.13
94	8	19	15	35 ^b	1	2	19	44			.11
95	14	33 ^b	1	2	27	63					-.03
96	9	21	28	65 ^b	4	9					-.05
97 ^a	2	5	12	28	9	21	20	47 ^b			-.13
98	7	16	8	19	6	14	22	51 ^b			.08
99 ^a	7	16	7	16	19	44	10	23 ^b			-.09
Task Test #120											
100 ^a	0	0	1	2	42	98 ^b	0	0			.37
101	2	5	37	86 ^b	2	5					.21
102	28	65 ^b	4	9	11	26					.24
103	9	21	19	44	14	33 ^b					.23
104	8	19	18	42	17	40 ^b					-.02
105	18	42	10	23 ^b	14	33					.17
106	12	28 ^b	13	30	17	40					.10
107	8	19	26	61 ^b	9	21					.25
108	14	33 ^b	18	42	11	26					.16
109	15	35	16	37 ^b	12	28					.08

APPENDIX B: (Continued)

Item	Response Alternative										r
	A		B		C		D		E		
	N	%	N	%	N	%	N	%	N	%	
Task Test #322											
110	13	30	2	5	21	49 ^b	7	16			.14
111	17	40 ^b	8	19	16	37	2	5			.20
112 ^a	24	56	1	2	11	26	7	16 ^b			.26
113	6	14	10	23	18	42 ^b	9	21			.34
114	9	21	2	5	30	70 ^b	2	5			.07
115	9	21	34	79 ^b	0	0					.00
116	0	0	3	7	30	70 ^b	10	23			.07
117 ^a	24	56	6	14 ^b	13	30					.20
118	18	42	5	12	20	47 ^b					.20
119	5	12	2	5	31	72 ^b	5	12			.10
120	2	5	7	16	12	28	22	51 ^b			.30
121	6	14	16	37 ^b	3	7	18	42			.10
Task Test #340											
123	3	7	23	54 ^b	3	7	14	33			.10
124	25	58 ^b	1	2	11	26	6	14			.13
125	18	42 ^b	6	14	13	30	6	14			.28
126	2	5	9	21	15	35 ^b	17	40			.03
127	12	28	16	37	14	33 ^b					.22
128	7	16 ^b	10	23	9	21	17	40			.23
Task Test #284											
129	6	14	13	30	23	54 ^b	1	2			.18
130 ^a	36	84 ^b	7	16							-.04
131	3	7	8	19	32	74 ^b					.03
132	0	0	3	7	5	12	35	81 ^b			.36
133	29	67 ^b	11	26	3	7					.05
Task Test #275											
134 ^a	0	0	42	98 ^b	1	2	0	0			.11
135	1	2	1	2	3	7	38	88 ^b			.24
136	1	2	27	63 ^b	6	14	9	21			.11
137 ^a	4	9 ^b	4	9	0	0	35	81			.26
138	4	9	35	81 ^b	3	7					.29
139	1	2	16	37 ^b	1	2	22	51	3	7	.31
Task Test #446											
140	19	44 ^b	24	56	0	0					.10
141 ^a	9	21	13	30	11	26 ^b	10	23			-.15
142	13	30	5	12	17	40 ^b	8	19			.11
143	1	2	14	33	2	5	25	58 ^b			-.09
144	2	5	20	47	13	30 ^b	8	19			.07

APPENDIX B: (Continued)

Item	Response Alternative										r
	A		B		C		D		E		
	N	%	N	%	N	%	N	%	N	%	
Task Test #503											
145	5	12	6	14	4	9	28	65 ^b			.17
146	2	5	10	23	1	2	19	44 ^b	11	26	-.03
147	26	61 ^b	13	30	1	2	3	7			.57
148	4	9	5	12	6	14	23	54 ^b	4	9	.14
149	10	23	22	51 ^b	10	23	1	2			.22
150	23	54 ^b	3	7	16	37	1	2	0	0	-.14
151 ^a	35	81	4	9 ^b	4	9					-.32
152	30	70 ^b	3	7	10	23					-.15
Task Test #155											
153 ^a	0	0	41	95 ^b	2	5					.11
154	0	0	2	5	39	91 ^b	2	5			.08
155	5	12	4	9	31	72 ^b	3	7			.21
156	7	16 ^b	18	42	10	23	8	19			-.11
157	31	72 ^b	1	2	11	26					.21
158 ^a	41	95 ^b	1	2	1	2	0	0			.11
159 ^a	42	98 ^b	0	0	0	0	1	2			.24
Task Test #162											
160 ^a	1	2	0	0	42	98 ^b	0	0			.19
161	0	0	1	2	4	9	38	88 ^b			.15
162	2	5	0	0	20	47 ^b	21	49			.12
163 ^a	0	0	42	98 ^b	1	2	0	0			.01
164	6	14	35	81 ^b	2	5	0	0			.13
165	1	2	40	93 ^b	2	5					.03
Task Test #421											
167 ^a	23	54	20	47 ^b	0	0					.22
168 ^a	4	9	1	2	38	88 ^b	0	0			-.10
169	6	14	2	5	24	56 ^b	11	26			.42
170	12	28	6	14	1	2	24	56 ^b			.29
171	0	0	27	63 ^b	12	28	4	9			-.10
172	6	14	22	51 ^b	13	30	2	5			.09
173 ^a	33	77 ^b	1	2	0	0	9	21			-.15
174	9	21	2	5	2	5	30	70 ^b			.03
Task Test #181											
175	1	2	7	16	4	9	31	72 ^b			-.22
176	4	9	10	23	8	19	21	49 ^b			-.22
Task Test #477											
177 ^a	2	5	1	2	1	2	39	91 ^b			-.07
178	4	9	1	2	38	88 ^b					-.07
179	0	0	33	77 ^b	9	21					-.10
180	2	5	4	9	19	44	18	42 ^b			-.36

APPENDIX B: (Concluded)

Item	Response Alternative										r
	A		B		C		D		E		
	N	%	N	%	N	%	N	%	N	%	
Task Test #477 (Continued)											
181 ^a	0	0	1	2	35	81 ^b	7	16			.07
182 ^a	0	0	36	84 ^b	7	16					-.24
183	37	86 ^b	2	5	2	5	2	5			.00
184	4	9	38	88 ^b	1	2					-.27
185	2	5	3	7	21	49	17	40 ^b			-.06
186 ^a	43	100 ^b	0	0	0	0	0	0			.00
187 ^a	2	5	40	93 ^b	1	2					.13
Task Test #549											
188	19	44	8	19	0	0	16	37 ^b			.10
189 ^a	0	0	0	0	42	98 ^b					.29
190	19	44 ^b	22	51	0	0	2	5			.20
191	41	95 ^b	1	2	1	2					.18
192 ^a	41	95 ^b	0	0	2	5					-.07
193 ^a	0	0	1	2	42	98 ^b					.29
194	42	98 ^b	0	0	1	2					.29
195 ^a	43	100 ^b	0	0	0	0					.00
Task Test #300											
196	14	33 ^b	10	23	7	16	12	28			.04
197	6	14	2	5	0	0	35	81 ^b			.33
198	0	0	27	63 ^b	15	35	1	2			-.04
199 ^a	17	40	3	7	21	49 ^b	2	5			.00
200	4	9	11	26	9	21	19	44 ^b			.11
201 ^a	40	93 ^b	1	2	0	0	2	5			.17
Task Test #238											
202	2	5	37	86 ^b	4	9	0	0			.25
203	15	35 ^b	5	12	12	28	11	26			.25
204	4	9	4	9	24	56 ^b	4	9	7	16	.26
205	13	30	28	65 ^b	2	5					.13
206	2	5	20	47 ^b	12	28	9	21			.16
207	8	19	6	14	26	61 ^b	3	7			.36
208 ^a	1	2	1	2	7	16	34	79 ^b			.08
209	10	23	17	40 ^b	12	28	4	9			.08
210	29	67 ^b	8	19	1	2	5	12			-.30
211	1	2	2	5	24	56 ^b	15	35			-.17
212	2	5	31	72 ^b	5	12	4	9			.41

Note: Percentages for items may not equal 100% due to missing data, multiple responses to an item, selection of a nonexistent alternative, or rounding.

^aItem deleted from the final version of the AGE job knowledge test.

^bIndicates the correct response.

^cIndicates the corrected item-total correlation.

APPENDIX C: AIRCREW LIFE SUPPORT PRETEST AND FINAL TEST CHARACTERISTICS

Characteristics	Pretest version					Final Version	
	I	MAC	TAC	SAC	Total	I	Total
# of Booklets	1	1	1	1	4	1	4
# of Task Tests	8	5	5	5	23	8	23
# of Items	70	52	68	73	263	65	258
% of Items with Illustrations	8%	19%	0%	0%	6%	9%	6%
Range of Items per Task Test	4-19	4-15	7-20	7-20	4-20	3-18	3-20
Distribution of Number of Items per Task Test	1-4 1-5 2-6 1-8 1-10 1-12 1-19	1-4 1-7 1-11 2-15	1-7 1-9 1-12 2-20	1-7 2-14 1-18 1-20	2-4 1-5 2-6 3-7 1-8 1-9 1-10 1-11 2-12 2-14 2-15 1-18 1-19 3-20	1-3 2-5 1-6 2-8 1-12 1-18	1-3 1-4 2-5 1-6 3-7 2-8 1-9 1-11 2-12 2-14 2-15 2-18 3-20
Average Number Items per Task	8.75	10.40	13.60	14.60	11.43	8.13	11.22
Range of Number of Choices per Item	2-5	3-5	3-5	3-5	2-5	2-4	2-5
Distribution of Choices	2-2 5-3 61-4 2-5	0-2 1-3 49-4 2-5	0-2 4-3 61-4 3-5	0-2 8-3 62-4 3-5	2-2 18-3 233-4 10-5	2-2 5-3 58-4	2-2 18-3 230-4 8-5
Printed Pages	28	19	23	23	93	26	91
Time to Finish Test (Range)	.5 to 1 hr.	.5 to 1 hr.	.5 to 1 hr.	.5 to 1 hr.	1 to 2 hrs.	.5 to 1 hr.	1 to 2 hrs.

Note. No description of the pilot test is reported due to the unstandardized nature of the pilot test.

**APPENDIX D: ITEM ANALYSIS FOR PRETEST VERSION OF THE AIRCREW
LIFE SUPPORT (AFS 122X0) JOB KNOWLEDGE TEST**

Phase I
(N = 36)

Item	Response Alternative										r ^d
	A		B		C		D		E		
	N	%	N	%	N	%	N	%	N	%	
Task Test #389											
1	29	81 ^c	6	17	0	0	1	3			-.21
2	5	14	6	17	14	39	11	31 ^c			.00
3	30	83 ^c	3	8	1	3	2	6			.24
4 ^a	0	0	36	100 ^c	0	0	0	0			.00
Task Test #199											
5	34	94 ^c	1	3	1	3	0	0			.17
6	28	78 ^c	0	0	8	22	0	0			-.09
7	6	17	28	78 ^c	1	3	1	3			.13
8	3	8	6	17	19	53 ^c	8	22			.10
9	1	3	1	3	1	3	33	92 ^c			.23
10	2	6	3	8	24	67 ^c	7	19			-.16
Task Test #315											
11	1	3	3	8	28	78 ^c	4	11			.44
12	6	17	30	83 ^c							.35
13	6	17	17	47	4	11	9	25 ^c			.34
14	5	14	24	67 ^c	5	14	2	6			.32
15 ^{a,b}	28	78 ^c	4	11	2	6	2	6			-.02
16	29	81 ^c	0	0	5	14	2	6			.36
Task Test #320											
17	1	3	25	69 ^c	4	11	6	17			.20
18	24	67 ^c	8	22	2	6	2	6			-.12
19	13	36 ^c	7	19	11	31	5	14			.08
20	1	3	11	31	20	56 ^c	4	11			-.05
21	14	39	8	22 ^c	2	6	12	33			.05
Task Test #383											
22	10	28	19	53 ^c	4	11	3	8			.42
23	16	44 ^c	11	31	9	25					.50
24	30	83 ^c	3	8	0	0	3	8			.25
25	11	31	21	58 ^c	4	11					.33
26	3	8	4	11	0	0	29	81 ^c			.52
27	2	6	11	31	23	64 ^c					.21
28	19	53	17	47 ^c							-.36
29	0	0	33	92 ^c	3	8					.31

APPENDIX D: (Continued)

Item	Response Alternative										r
	A		B		C		D		E		
	N	%	N	%	N	%	N	%	N	%	
Task Test #330											
30	32	89 ^c	0	0	4	11	0	0			.06
31	3	8	5	14	26	72 ^c	2	6			.55
32	0	0	32	89 ^c	0	0	4	11			.24
33	13	36	0	0	22	61 ^c	1	3			.17
34	5	14	0	0	27	75 ^c	4	11			-.01
35	21	58 ^c	3	8	11	31	1	3	0	0	-.02
36	0	0	2	6	1	3	1	3	32	89 ^c	.10
37	26	72 ^c	10	28	0	0	0	0			.24
38	9	25	17	47 ^c	3	8	7	19			.18
39	26	72 ^c	7	19	1	3	2	6			.39
40	31	86 ^c	3	8	0	0	2	6			.23
41	1	3	0	0	5	14	30	83 ^c			.44
Task Test #380											
42	1	3	10	28 ^c	7	19	18	50			.38
43	7	19	1	3	1	3	27	75 ^c			.45
44	20	56 ^c	11	31	5	14	0	0			.41
45	0	0	1	3	24	67 ^c	11	31			.16
46	8	22	18	50 ^c	7	19	2	6			.03
47	23	64 ^c	8	22	5	14	0	0			.24
48	6	17	16	44 ^c	3	8	11	31			-.03
49	4	11	20	56	11	31 ^c	1	3			.28
50	1	3	10	28	5	14 ^c	20	56			-.01
51	4	11	5	14	25	69 ^c	2	6			.28
52	2	6	4	11	29	81 ^c	1	3			.29
53	3	8	11	31 ^c	8	22	14	39			.40
54	0	0	22	61 ^c	7	19	7	19			.37
55	12	33 ^c	14	39	1	3	8	22			.24
56	6	17	20	56 ^c	10	28					.10
57 ^a	0	0	1	3	1	3	34	94 ^c			.29
58	5	14	22	61 ^c	2	6	7	19			.25
59	20	56 ^c	4	11	9	25	3	8			.60
60	26	72 ^c	4	11	3	8	3	8			.41

APPENDIX D: (Concluded)

Item	Response Alternative										r
	A		B		C		D		E		
	N	%	N	%	N	%	N	%	N	%	
	Task Test #295										
61	1	3	28	78 ^c	1	3	6	17			.07
62	15	42	1	3	1	3	19	53 ^c			-.10
63 ^{a,b}	0	0	1	3	30	83 ^c	5	14			.16
64	23	64 ^c	9	25	3	8	1	3			.09
65	8	22	16	44	7	19	5	14 ^c			.27
66	14	39	13	36 ^c	8	22	1	3			.04
67	12	33 ^c	4	11	7	19	13	36			-.19
68	14	39 ^c	5	14	11	31	6	17			.35
69	0	0	16	44 ^c	3	8	17	47			.44
70 ^a	2	6	33	92 ^c	0	0	0	0			-.04

Note: Percentages for items may not equal 100% due to missing data, multiple responses to an item, selection of a nonexistent alternative, or rounding.

^aItem deleted from final version of Aircrew Life Support Job Knowledge Test.

^bDuplicate item.

^cIndicates the correct response.

^dIndicates the corrected item-total correlation.

**APPENDIX E: ITEM ANALYSIS FOR PRETEST VERSION OF
THE AIRCREW LIFE SUPPORT (AFS 122X0)
MAC, SAC, AND TAC JOB KNOWLEDGE TEST**

Phase II MAC						
(N = 14)						
Response Alternative						
Item	A	B	C	D	E	r ^b
Task Test #346						
71	3	6 ^a	3	2		.27
72	7 ^a	0	5	2		.44
73	4 ^a	0	6	4		.29
74	2	3	9 ^a	0		.33
75	1	1	1	11 ^a		.66
76	5	1	6 ^a	2		.46
77	8	4 ^a	1	1		.29
78	0	1	1 ^a	12		.10
79	0	1	3	10 ^a		.72
80	1	8 ^a	0	5		.19
81	1	0	7	6 ^a		.66
Task Test #379						
82	0	2 ^a	1	11		.04
83	4	8 ^a	2	0		.16
84	7	3	2	2 ^a		-.21
85	11 ^a	1	1	1		.40
86	6	6 ^a	2	0		-.22
87	0	1	1	12 ^a		-.09
88	1	6 ^a	0	7		-.30
89	4 ^a	3	2	5		.36
90	6	3	4 ^a	1		-.27
91	1	2	1	10 ^a		-.60
92	2	0	12 ^a	0		.48
93	12 ^a	2	0	0		.18
94	5 ^a	6	3	0	0	.03
95	3 ^a	0	6	5		.13
96	8 ^a	1	2	3		.06
Task Test #382						
97	3	5 ^a	5	1		.16
98	3	0	2	9 ^a		-.05
99	6	5 ^a	1	2		.45
100	2	0	12 ^a	0		.36
101	5	5	2	2 ^a		-.11
102	2 ^a	0	12	0		.34
103	5	8 ^a	1			.29
104	1	3	8 ^a	2		.68
105	8	1	5 ^a	0		.70
106	0 ^a	2	1	11		.00
107	1	0	12	1 ^a		.22

APPENDIX E: (Continued)

Phase II MAC (Continued)

(N = 14)

Response Alternative

Item	A	B	C	D	E	r
Task Test #382 (Continued)						
108	7	7	0	0 ^a		.00
109	5	4 ^a	2	3		.53
110	0	0	1	13 ^a		.16
111	0	0	14 ^a	0		.00
Task Test #210						
112	0 ^a	0	10	2	2	.00
113	2	4	7 ^a	1		.00
114	5	5 ^a	0	4		.11
115	2	8	1	3 ^a		-.13
Task Test #224						
116	9 ^a	0	3	2		.74
117	0	0	1	13 ^a		.28
118	1	11 ^a	1	1		.16
119	0	0	0	14 ^a		.00
120	0	13 ^a	1	0		.28
121	5 ^a	6	1	2		.29
122	4	3	7a	0		-.06

Phase II SAC

(N = 15)

Task Test #303

71	4	5	5 ^a	1		.00
72	0	11 ^a	2	2		.45
73	0	3 ^a	1	2	8	.28
74	1	0	11 ^a	3		.61
75	0	14 ^a	1	0		.44
76	0	7	4	4 ^a		.44
77	14 ^a	0	1	0		-.15

Task Test #349

78	15 ^a	0	0			.00
79	0	1	14 ^a	0		.31
80	3	11 ^a	1	0		.46
81	14	0	0	1 ^a		.29
82	14 ^a	0	0	1		.69
83	0	1	0	14 ^a		-.15
84	12 ^a	0	3	0		.10
85	0	0	15 ^a	0		.00
86	2	11 ^a	0	2		.10
87	0	0	15 ^a	0		.00
88	15 ^a	0	0	0		.00
89	2	0	9	4		.25

APPENDIX E: (Continued)

Phase II SAC (Continued)						
(N = 14)						
Response Alternative						
Item	A	B	C	D	E	r
Task Test #349 (Continued)						
90	15 ^a	0	0	0		.00
91	11 ^a	3	1			.17
92	0	0	3	12 ^a		.68
93	1	0	10	4 ^a		.25
94	0	4	7 ^a	3		.18
95	2	13 ^a	0	0		.80
96	10	1	0	4 ^a		.11
97	2	5	1	7 ^a		-.16
Task Test #398						
98	0	12 ^a	1	2	0	-.18
99	0	11 ^a	1	0	3	-.38
100	2	2	9 ^a	2		-.03
101	1	2	0	12 ^a		-.08
102	0	0	15 ^a	0		.00
103	7 ^a	3	5	0		.51
104	2	0	13 ^a	0		.25
105	0	14 ^a	1	0		-.01
106	11 ^a	3	0	1		.25
107	1	11 ^a	3	0		.37
108	0	0	6	9 ^a		.05
109	1	3	5 ^a	6		-.57
110	11 ^a	2	2			.37
111	0	5	8 ^a	2		.21
Task Test #359						
112	12 ^a	2	1			.50
113	9 ^a	3	3	0		.60
114	1	1	0	13 ^a		.28
115	1	4	10 ^a			.55
116	0	12 ^a	2	1		.45
117	1	5	8	1 ^a		.25
118	6	7 ^a	2			.03
119	0	0	8 ^a	6		.62
120	9 ^a	2	1	3		.79
121	6	5 ^a	0	4		.11
122	1	1	2	11 ^a		.67
123	9 ^a	4	2			.51
124	4	9 ^a	1	1		-.03
125	1	9	5 ^a	0		-.21
126	4	3	8 ^a	0		.19
127	1	13 ^a	0	1		.05
128	4	1	1	9 ^a		.89
129	2	1	12 ^a	0		-.19

APPENDIX E: (Continued)

Phase II SAC (Continued)

(N = 14)

Item	Response Alternative					r
	A	B	C	D	E	
Task Test #310						
130	10 ^a	1	3	1		-.19
131	1	0	0	14 ^a		.46
132	0	9 ^a	6	0		-.42
133	15 ^a	0	0	0		.00
134	0	0	4 ^a	11		.11
135	0	6 ^a	9			.08
136	1	4	9 ^a	1		.00
137	6 ^a	5	1	3		.08
138	12	1 ^a	2	0		.15
139	2	0	12 ^a	1		.05
140	6 ^a	5	3	1		-.20
141	1 ^a	1	0	13		-.38
142	9	5 ^a	1	0		.14
143	1	11 ^a	1	1		.13

Phase II TAC

(N = 7)

Task Test #303						
71	2	1	3 ^a	1		-.26
72	0	3 ^a	2	2		-.26
73	0	3 ^a	1	0	3	-.75
74	2	0	5 ^a	0		.09
75	0	6 ^a	0	1		-.71
76	0	3	1	3 ^a		-.26
77	7 ^a	0	0	0		.00

Task Test #349						
78	7 ^a	0	0			.00
79	0	0	7 ^a	0		.00
80	2	5 ^a	0	0		-.52
81	7	0	0	0 ^a		.00
82	7 ^a	0	0	0		.00
83	0	0	0	7 ^a		.00
84	6 ^a	0	1	0		.11
85	0	0	7 ^a	0		.00
86	2	4 ^a	0	1		.23
87	0	0	6 ^a	1		.11
88	7 ^a	0	0	0		.00
89	0	0	6 ^a	1		-.55
90	7 ^a	0	0	0		.00
91	7 ^a	0	0			.00
92	0	1	6	0 ^a		.11
93	0	0	5	2 ^a		.49

APPENDIX E: (Continued)

Phase II TAC (Continued)						
(N = 7)						
Response Alternative						
Item	A	B	C	D	E	r
Task Test #349 (Continued)						
94	1	3	3 ^a	0		.28
95	0	7 ^a	0	0		.00
96	0	5	0	2 ^a		-.44
97	0	2	2	3 ^a		-.48
Task Test #311						
98	6 ^a	1	0	0		.25
99	0	0	7 ^a			.00
100	0	0	0	0	7 ^a	.00
101	0	1	2	4 ^a		.24
102	4 ^a	0	3	0		.94
103	0	6 ^a	0	1		.46
104	3	1	2 ^a	1		.19
105	0	4	3 ^a	0		-.02
106	2 ^a	4	1	0		-.10
107	1	1	3	2 ^a		-.24
108	2	0	1	4 ^a		.94
109	4	3 ^a	0	0		.77
Task Test #278						
110	0	0	0	7 ^a		.00
111	7 ^a	0	0			.00
112	0	0	7 ^a	0		.00
113	6 ^a	1	0	0		.17
114	1	1	5 ^a	0		-.06
115	0	7 ^a	0	0		.00
116	0	0	0	7 ^a		.00
117	1	0	2	4 ^a		.42
118	0	7 ^a	0	0		.00
Task Test #483						
119	1	2 ^a	3	1		-.68
120	1	0	1	5 ^a		.72
121	3	0	4 ^a	0		.33
122	0	0	0	7 ^a		.00
123	2	2 ^a	2	1		.47
124	0	1	6 ^a	0		.49
125	4 ^a	0	0	3		.33
126	0	0	0	6 ^a	1	-.34
127	1	0	6 ^a	0		.24
128	3 ^a	2	2	0		.06
129	7 ^a	0	0	0		.00
130	1	1	1	4 ^a		.93
131	0	4 ^a	0	3		.93

APPENDIX E: (Concluded)

Phase II TAC (Concluded)						
(N = 7)						
Response Alternative						
Item	A	B	C	D	E	r
Task Test #483(Continued)						
132	0	1	1	5 ^a		.62
133	0	1	6 ^a	0		.49
134	3	0	4 ^a	0		.24
135	1 ^a	1	3	2		.40
136	5 ^a	0	0	2		.12
137	1	1	4 ^a	1		.16
138	0	0	5 ^a	2		.72

Note: Mean percent correct is not reported for Phase II due to the small sample sizes.

^aIndicates the correct response.

^bIndicates the corrected item-total correlation.

APPENDIX F: PERSONNEL JKT CHARACTERISTICS

Characteristics	Pretest and Final Test						Total
	I	C&T	Records	Manning	Outbound	Separations	
# of Booklets	1	1	1	1	1	1	6
# of Task Tests	8	10	10	10	10	10	58
# of Items	47	53	41	47	58	55	301
% of Items with Illustrations	23%	57%	22%	34%	86%	60%	50%
Range of Items per Task Test	2-10	4-7	2-8	2-8	3-14	2-13	2-14
Distribution of Number of Items per Task Test	1-2	2-4	1-2	2-2	3-3	1-2	5-2
	2-3	4-5	4-3	1-3	2-4	1-3	11-3
	2-5	3-6	1-4	1-4	2-5	3-4	9-4
	1-9	1-7	3-5	3-5	1-6	2-5	16-5
	2-10		1-8	1-6	1-11	1-7	5-6
				1-7	1-14	1-8	3-7
				1-8		1-13	3-8
							1-9
							2-10
Average Number Items per Task	5.87	5.30	4.10	4.7	5.8	5.5	5.19
Range of Number of Choices per Item	3-5	3-4	3-5	3-5	2-5	2-5	2-5
Distribution of Choices	2-3	1-3	2-3	8-3	11-2	4-2	15-2
	42-4	52-4	38-4	35-4	11-3	7-3	31-3
	3-5		1-5	4-5	34-4	42-4	243-4
					2-5	2-5	12-5
Printed Pages	19	32	26	28	37	30	172
Minutes to Finish							
Test (Range)	15-38	25-33	22-38	32-39	22-39	15-37	33-77

Note: No description of the pilot test is reported due to the unstandardized nature of the pilot test. Pretest and final test information is reported as one because no revisions were made to the Personnel Phase I due to small sample sizes.

**APPENDIX G: ITEM ANALYSIS FOR PRETEST VERSION OF THE
PHASE I PERSONNEL SPECIALTY (AFS 732X0) JOB KNOWLEDGE TEST**

Item	Response Alternative										r ^b
	A		B		C		D		E		
	N	%	N	%	N	%	N	%	N	%	
Task Test 733											
1	19	76 ^a	1	4	5	20					.46
2	0	0	2	8	5	20	18	72 ^a			.36
3	1	4	14	56	0	0	10	40 ^a			.17
4	0	0	1	4	13	52 ^a	11	44			.24
5	16	64	6	24 ^a	2	8	1	4			.16
Task Test 121											
6	0	0	0	0	22	88 ^a	3	12			.21
7	11	44	5	20 ^a	7	28	2	8			.21 [*]
Task Test 131											
8	0	0	3	12	0	0	21	84 ^a	1	4	.53
9	10	40	1	4	0	0	14	56 ^a			.47
10	22	88 ^a	0	0	0	0	3	12			.54
11	0	0	23	92 ^a	2	8					.45
12	0	0	1	4	23	92 ^a	1	4			.45
Task Test 35											
13	0	0	1	4	0	0	23	92 ^a			.37
14	0	0	25	100 ^a	0	0	0	0			.71
15	2	8	22	88 ^a	1	4					.37
16	3	12	18	72 ^a	3	12	1	4			.43
17	0	0	1	4	24	96 ^a					.74
18	0	0	1	4	1	4	23	92 ^a			.67
19	1	4	0	0	22	88 ^a	2	8			.50
20	14	56 ^a	7	28	0	0	4	16			.34
21	2	8	1	4	22	88 ^a					.50
Task Test 896											
22	3	12	7	28	14	56 ^a	1	4			.00
23	3	12	16	64 ^a	3	12	2	8			.38
24	5	20	4	16	10	40 ^a	6	24			.27
Task Test 140											
25	22	88 ^a	2	8	0	0	1	4			.54
26	4	16	3	12	0	0	3	12	15	60 ^a	.36
27	3	12	4	16	9	36 ^a	9	36			.24

APPENDIX G: (Concluded)

Item	Response Alternative										r
	A		B		C		D		E		
	N	%	N	%	N	%	N	%	N	%	
Task Test 719											
28	0	0	1	4	0	0	24	96 ^a			.91
29	25	100 ^a	0	0							.74
30	1	4	0	0	0	0	24	96 ^a			.91
31	0	0	19	76 ^a	6	24					.51
32	25	100 ^a	0	0	0	0					.74
33	0	0	1	4	22	88 ^a	2	8			.69
34	0	0	23	92 ^a	2	8					.70
35	0	0	0	0	2	8	23	92 ^a			.45
36	24	96 ^a	1	4							.91
37	1	4	0	0	3	12	21	84 ^a			.53
Task Test 116											
38	2	8	23	92 ^a	0	0					.61
39	0	0	24	96 ^a	1	4					.35
40	4	16	0	0	21	84 ^a					.23
41	2	8	1	4	1	4	21	84 ^a			.40
42	17	68 ^a	2	8	4	16	2	8			.44
43	6	24	7	28	1	4	10	40 ^a			.43
44	0	0	24	96 ^a	0	0	1	4			.60
45	14	56 ^a	7	28	0	0	4	16			.30
46	19	76 ^a	4	16	2	8					.00
47	1	4	7	28	0	0	17	68 ^a			.21

Note: Percentages for items may not equal 100% due to missing data, multiple response to an item, selection of a nonexistent alternative, or rounding.

^aIndicates the correct response.

^bIndicates the corrected item-total correlation.